Local ordering and dynamics of plastic crystals

Andrés Henao Aristizabal
Departament de Física
Universitat Politècnica de Catalunya - Barcelona Tech

A thesis submitted for the degree of
Doctor of Philosophy

Barcelona, July 2016

Supervised by

Elvira Guàrdia
Luis Carlos Pardo

Thesis supervised by

Elvira Guardia Manuel
Universitat Politècnica de Catalunya, Barcelona, Spain
Grup de Simulació per Ordinador en Matèria Condensada - SIMCON

Luis Carlos Pardo Soto
Universitat Politècnica de Catalunya, Barcelona, Spain
Grup de Caracterització de Materials - GCM

The research in this PhD thesis has been carried out within the doctoral programme of Computational and Applied Physics of the Department of the Universitat Politècnica de Catalunya - Barcelona Tech. It has been mainly supported by the predoctoral researcher FPU/UPC fellowship reference 43/721 of 26/05/11 and partly by the Spanish Ministry MINECO grants No. FIS2014-54734-P and FIS2012-39443-C02-01, the Government of Catalonia grants No. 2009SGR-1003 and 2014SGR-00581 and the fellowship from Colciencias "Doctorados en el exterior 2013".

Local ordering and dynamics of plastic crystals

Andrés Henao Aristizábal
Departament de Física, Universitat Politècnica de Catalunya

Abstract

Many materials of interest show a disordered phase, with their molecules presenting positional disorder, orientational disorder or both. Processes of interest to life also occur in these kind of phases as it is the case of liquid water present in many biochemical problems of interest such as protein folding. The study of these disordered phases is inherently difficult due to their lack of periodicity as in ordered crystals.

In this work, molecular dynamics simulations (MD) are used to study two disordered phases: the liquid and the plastic crystal. The configurations obtained by MD are analyzed using different methodologies to obtain insights into the local ordering of the molecules and into their dynamics. The local ordering of the systems is studied using an Euler angles method considering the molecules as rigid bodies. Probability distributions describing the positional and orientational molecular ordering are obtained. These probability distributions are analysed using information theory concepts that provide a deeper understanding of the results.

Two systems of interest are studied: hexachloroethane and water. The structure of the liquid and plastic crystal phases of hexacloroethane are investigated and the results from MD and the local ordering analyses are compared to previous works using simulations and neutron diffraction experiments. A Bayesian approach is used in order to obtain the intra-molecular structure of the molecule from the neutron diffraction experimental data, a set of parameters in better agreement with the plastic phase is obtained compared to previous electron diffraction experiments.

Liquid water at room conditions is studied on the short range order, presenting a contribution to the ongoing topic of its local structure. The ordering of water at distances beyond the first hydration shell is also studied using the information theory approach. Findings of this work support recent experimental and simulation data showing the existance of long range correlations in water. Finally, plastic crystals phases of water found at high pressures by MD are studied and compared to a high pressure liquid and the ice vii form by comparing their local ordering and dynamics.

The methodologies used in this work provide a powerful technique to analyse disordered phases and the results obtained will contribute to the understanding on the structure of such phases.

Acknowledgement

Thanks must go to my supervisors, Dr. Luis Carlos Pardo and Dra. Elvira Guàrdia Manuel, for being true to the guidance and motivation during all the process of this thesis. They were not only great supervisors, providing me with the resources needed and opportunities through the whole project, but they also offered me their friendship and they can always count on mine too. I have learned from them the commitment to scientific quality.

I'd like to thank Dra. Sylvia E. McLain at Oxford University and Dr. Christian D. Lorenz at King's College London for the fruitful stays in their groups, I have learned a lot in a short time. I give special thanks to Dr. Manel Canales for the introduction to the use of Gromacs, that was the main tool of this thesis.

The research task was made easy thanks to the administrative and technical support of Cristina, Ana, Montserrat and Alfredo, they do such a good job. Also in the doctoral office to Ana and Alicia. I give my sincere thanks to all the people at the SIMCON and GCM research groups: Josep Lluis Tamarit always gave me an encouraging smile, Grigori offered me his friendship and made easy my adaptation to the group, I spent so many good times away from research with the group to stay sane, Maria, Roberto, Michela, Gloria and many others. To the doctoral students within the group also thank you, Guillem, Ausias, Jordi, Amani, Jin, Efi, Alessandro, Pragya, Manesh and Guada, this journey was easier shared, and to Luis and Omar, it is possible!.

In the personal I am in debt with so many people that I can not put them all in here. Thanks to my second families in Tarragona and Barcelona, to my new friends in Barcelona, Nicaragua and Brasil. To David and Jorge Mario, who know what it is to leave home to follow the PhD dream, thank you for having your hand always ready to help. Thanks to the Ferreira-Torres family, they believed in me at the beginning of my professional career and also made this possible.

I owe what I am to my parents, it has been so hard to be away, but here it is the reward. Thanks to my sister, brother, nephews, family in law, mother in law: you have always been there for me. To all my friends and family in Colombia for making me feel home from far.

I finish giving my thanks to the most important person in this journey. Daniela, you have been the first witness of all my process to get here, the good, the fun, the sad, the bad, but you were always there for me and it was easier every day because I had a home where I wanted to be with you. I love you and all I can say is that I am always there for you.

List of Papers

I. **Insights into the determination of molecular structure from diffraction data using a Bayesian algorithm**
A. Henao, M. Rovira-Esteva, A. Vispa, J. Ll. Tamarit, E. Guardia and L. C. Pardo
Journal of Physics: Condensed Matter **25**(45), 454217 (2013)

II. **Characterizing ordering in liquids: An information theoretic approach**
L. C. Pardo, A. Henao and A. Vispa
Journal of Non-Crystalline Solids **407**, 220-227 (2015)

III. **Competing structures within the first shell of liquid C_2Cl_6: a molecular dynamics study**
A. Henao, S. Pothoczki, M. Canales, E. Guardia and L. C. Pardo
Journal of Molecular Liquids **190**, 121-125 (2014)

IV. **A continuous mixture of two different dimers in liquid water**
L. C. Pardo, A. Henao, S. Busch, E. Guardia and J. Ll. Tamarit
Physical Chemistry Chemical Physics **16**(44), 24479-24483 (2014)

V. **The structure of liquid water beyond the first hydration shell**
A. Henao, S. Busch, E. Guardia, J. Ll. Tamarit and L.C. Pardo
Physical Chemistry Chemical Physics **DOI: 10.1039/c6cp00720a**, (2016)

Role of the author in this work

This thesis is devoted to the presentation, analysis and discussion of results from computer simulations of disordered phases: liquid and plastic crystal. The results were all produced using the molecular dynamics (MD) package named GroMacS (Groningen Machine Simulation).

The MD simulations were all performed by the author using the computational resources of the Condensed Matter Simulation research group at Universitat Politècnica de Catalunya.

Papers I and II present the Information theory method to analyse the Euler angles and the Bayesian fit method used for the study of the simulations performed. They are contained in Sections 2.2, 2.3 and Chapter 3. The author performed the simulations used in those papers and contributed to the discussion and writing of the manuscript.

The original research presented in this thesis is contained in Chapters 4-7. Chapter 4 contains the fitting of the intramolecular structure from neutron diffractions in paper II, that analysis was performed by the author. Paper III contains all the results for the liquid phase of Hexachloroethane in Chapter 4, the author performed the simulations and wrote the manuscript with revision and discussion of the coauthors. The plastic phase results presented in Chapter 4 have not been published yet and are the work of the author. Chapter 5 is based on paper IV, the author performed the simulations and assisted in the analysis and writing of the manuscript. Chapter 6 is based on paper V, the author performed the simulation and the analyses were done in collaboration between the author and coauthors as well as the writing of the manuscript. Finally, Chapter 7 presents results that have not been published yet. They are the complete work of the author under the supervision of Dr. Elvira Guàrdia and Dr. Luis Carlos Pardo.

Contents

Chapter 1

Introduction

1.1 About disordered phases

There are two ideal limits of ordered and disordered phases: the perfect crystal and the ideal gas, respectively. In a perfect crystal the intermolecular interactions are strong and the positions and orientations are fully correlated, that makes possible to treat the displacements as small perturbations of a fixed equilibrium structure. Because of this, the structure of a solid can be known from ordering rules, i.e knowing the unit lattice. In the opposite side, in an ideal gas the intermolecular interactions are negligible, so they can be treated as small perturbations, therefore the positions and orientations are completely uncorrelated.

Liquids are in an intermediate state between perfect crystals and ideal gases. Hence, the study of the structure in liquids is a difficult problem. In liquids, the intermolecular interactions are strong enough and they cannot be treated as perturbations as in gases. At the same time the molecular displacements are large, so they cannot be treated as small perturbations from an ideal lattice as in solids.

In practice, in nature there are not perfectly disordered phases such as the ideal gas, neither a perfectly ordered structure. There are vibrations at finite temperatures and even at the zero point, the uncertainty principle

does not allow particles to stay still (27, 42). Most materials lie between the ordered and disordered limits. There are other states of matter showing disorder in their positional or orientational degrees of freedom. Plastic crystals are one of those phases, where the center of mass of molecules form a regular lattice as in solids, but the molecules are dynamically disordered with respect to their orientational degrees of freedom as in liquids. Figure 1.1 shows a scheme on the different degrees of freedom in liquids, plastic crystals and ordered crystals. The plastic crystal phase is interesting from the point of view of physics because it can be used to understand the liquid-glass transition, when cooling a viscous liquid the degrees of freedom can be frozen in a disordered state called a glass. Most plastic crystals are formed by molecules more or less globular shaped, which favours their reorientation. Given the translational symmetry in the plastic crystals, they are much simpler to treat theoretically compared to the glasses (6).

The complete picture of the structure of liquids is still not clear (36). This has been a matter of discussion since long time ago. In 1962 Bernal (4) introduced the concept of hard spheres packing for the structure of simple liquids on his Bakerian lecture. A later review by Chandler in 1978 (7) underlined the importance of the idea, originated with the work of Van der Waals long ago, that repulsive forces dominate the liquid structure. Following this idea, the shape of molecules would determine the intermolecular correlations. But this is not always the case: Berthier and Tarjus addressed the role of attractive forces for the viscous regime of a liquid (5)). The problem of locally ordered packing is key for the stability of liquids and it has been discussed in metallic liquids and glasses (26). Recently Ingebritgsten et.al (17) discussed the definition of a (Roskilde) simple liquid regarding whether or not the long-range interactions can be ignored and quantified this via the virial potential-energy correlation. Computational techniques have become an essential tool in condensed matter research: molecular dynamics (MD) (1), Reverse Monte Carlo (RMC) (15, 24) and Empirical Potential Structure Refinement(EPSR) (32, 33) are some of the widely used methods to study disordered systems. These, along with the experimental data (Light scattering, dielectric constant measurements, neutron and X-

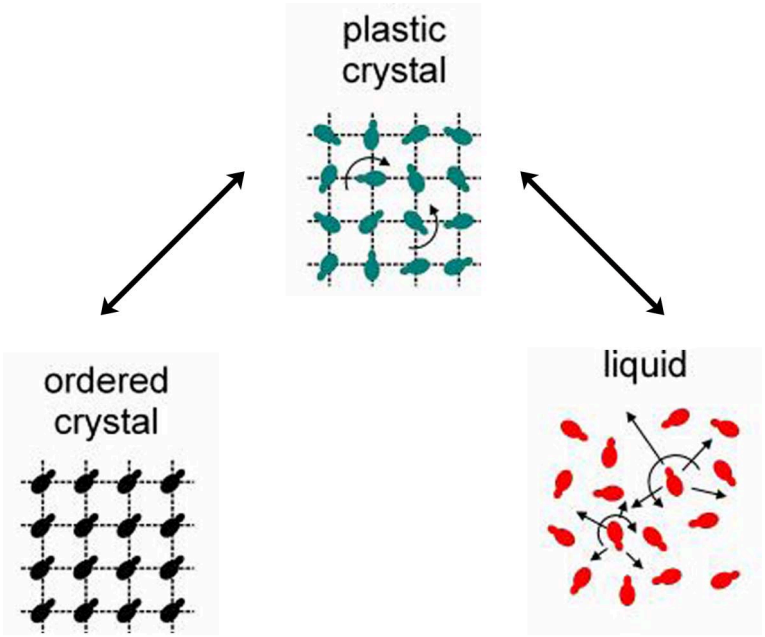

Figure 1.1: Scheme on the different degrees of freedom for a perfect crystal, a plastic crystal and a liquid. In a perfect crystal positions and orientations are ordered. In a liquid, both positions and orientations are disordered. The plastic crystal is an intermediate phase between the liquid and the crystal: the positions are ordered as in a crystal, but the orientations are disordered as in a liquid. *Modified from Ref. (6)*

ray scattering), are the common tools used to unravel the principles behind the ordering of these structures. This thesis uses the MD method to study disordered phases and comparison to experiments is done when available.

There is a lack of a universal theory to describe liquids and how they fall out of equilibrium forming a glass, but what seems to be sure is that the liquid structure is suspected to play an important role (22). In any case there is not a universal way to characterize the structure of a liquid and many different approaches have been used in the past (8, 9, 10, 13, 20, 21, 28, 30, 31, 35, 38, 40, 41). The general purpose of this thesis is to contribute to the understanding on the structure of disordered phases, namely liquids and plastic crystals.

1.2 Water

Water plays a central role in life. Not only it is present in about 70% of the surface of the earth, but most of the biological processes that are fundamental to life happen in the presence of water. In this thesis, the local structure of water at room conditions and also its structure on the long range will be investigated. The dynamics and local ordering of two new phases of water: plastic crystals at high pressure (2, 3, 39) is also studied.

The local structure of liquid water is still a controversial subject. In contradiction to the classical picture of a symmetric tetrahedral local order of neighboring water molecules around a central one (16, 18, 34, 37), two main alternative descriptions have been proposed: a local order with only one acceptor and one donor that would lead to a chain and ring structure of liquid water (13) and a two-state model that would imply an inhomogeneous picture of the water structure (11, 14, 25). Kühne and Khaliullin (18) proposed a way to reconcile the classical picture of the short range order of water with its counterparts by means of an asymmetry of the electronic interaction between molecular contacts. This asymmetry was found in the interaction between the two strongest hydrogen bond donors and acceptors. Later, they also found a difference in the geometry of these hydrogen bonds (19).

Concerning the structure beyond the first hydration shells of water, little is known. Recent results from second-harmonic light experiments revealed long-range molecular orientation correlations in liquid water (29). Liu et al.(23) also found angular correlations persisting on a length scale of $\sim$40 Å using Density Functional Theory. In that work it seemed that the special molecular arrangements are not correlated to the local density probed by the oxygen-oxygen partial radial distribution function $g_{OO}(r)$. Higo et al. also found angular correlations in bulk water (12). They obtained a dipole chain order for distances up to $\sim$10 Å.

Finally, it has been recently found two possible plastic phases of water by means of MD simulations (2, 3, 39). The diagram phase region where

the new plastic phases have been found are yet unaccessible experimentally, they are in the region of 8 GPa in pressure. If found experimentally, it would mean that the computational models for water are pointing in the right direction, the opposite would mean that the model is too spherical and that it should be reconsidered. Either way, it represents one step forward for computational models of water and for the understanding of this important substance.

1.3 Structure of the thesis

The remainder of this thesis is structured as follows. Firstly, in Chapter 2, the computational methods used are introduced: the molecular dynamics simulation (MD) method is described highlighting its most important concepts and the main aspects included in a typical simulation, next, the Euler angles analysis is described, which has been a fundamental analysis tool to study the structure for the systems simulated by MD and also a bayesian fit analysis is introduced, it will be used in Chapter 4 to obtain the intramolecular structure from neutron diffraction experiments.

Chapter 3 describes the information theory, giving an overview of the most important concepts and showing its application to the study of disordered phases. Although this theory has been used before to characterize the structure of liquids, it will be used in later chapters to obtain novel results on the structure of water. At the end of the chapter an example is used to present its application in the context of this thesis.

The original research presented in this thesis is split into 4 distinct subprojects, each of which are given their own chapter (Chapter 4-7). The subprojects are united by a common theme: each is an investigation, using MD simulations as the main tool, of the short range order and some dynamical aspects of two disordered phases, liquids and plastic crystals (or orientationally disordered crystals). In Chapter 4, the liquid and plastic phases of Hexachloroethane (C_2Cl_6) are studied and their short range structure is compared. The MD results are compared to previous neutron diffraction

experiments of the plastic phase and new insights to their structure is presented.

The second and third sub-projects (Chapters 5 and 6) deal with the structure of liquid water at ambient conditions. Chapter 5 is a study on the local structure of liquid water, which is still a controversial subject. Chapter 6 presents an investigation on the structure of water beyond the first hydration shell using the information theory approach. Finally, the fourth subproject (Chapter 7) also deals with water, but now in the high pressure range (8 GPa) where two new plastic phases have been found by means of MD simulations. The structure and dynamics of these new plastic phases and also those of a high pressure liquid and the ice vii crystal are studied.

And last, in Chapter 8 the findings of all the sub-projects are summarised, and an overview about interesting studies that might be carried out in the future are presented.

1.4 Bibliography

[1] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. i. general method. *J. Chem. Phys.*, 31:459–466, 1959. 2

[2] J. Aragones and C. Vega. Plastic crystal phases of simple water models. *The Journal of chemical physics*, 130(24):244504, 2009. 4

[3] J. Aragones, M. Conde, E. Noya, and C. Vega. The phase diagram of water at high pressures as obtained by computer simulations of the tip4p/2005 model: the appearance of a plastic crystal phase. *Physical Chemistry Chemical Physics*, 11(3):543–555, 2009. 4

[4] J. D. Bernal. The bakerian lecture, 1962. the structure of liquids. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences.*, 280(1382):299–322, July 1964. 2

[5] L. Berthier and G. Tarjus. Nonperturbative effect of attractive forces in viscous liquids. *Physical Review Letters.*, 103:170601, October 2009. 2

[6] R. Brand, P. Lunkenheimer, and A. Loidl. Relaxation dynamics in plastic crystals. *The Journal of chemical physics*, 116(23):10386–10401, 2002. 2, 3

[7] D. Chandler. Structures of molecular liquids. *Annual Review of Physical Chemistry.*, 29:441–471, 1978. 2

[8] M. Chaplin. A proposal for the structuring of water. *Biophysical chemistry*, 83(3):211–221, 2000. 3

[9] A. De Santis and D. Rocca. The local order in liquid water studied through restricted averages of the angular correlation function. *The Journal of chemical physics*, 107(22):9559–9568, 1997. 3

[10] A. De Santis and D. Rocca. Angular distribution functions and specific local structures in liquid water. *The Journal of chemical physics*, 107 (23):10096–10101, 1997. 3

[11] C. H. *et al. PNAS*, 106:15214–15218, 2009. 4

[12] J. H. *et al. PNAS*, 98:5961–5964, 2001. 4

[13] P. W. *et al. Science*, 304:995–999, 2004. 3, 4

[14] T. T. *et al. Chem. Phys. Lett.*, 460:387–400, 2008. 4

[15] O. Gereben, P. Jovari, L. Temleitner, and L. Pusztai. A new version of the rmc++ reverse monte carlo programme, aimed at investigating the structure of covalent glasses. *J. Optoel. Adv. Mat.*, 9(6):3021, 2007. 2

[16] T. Head-Gordon and G. Hura. *Chem. Rev.*, 102:2651, 2002. 4

[17] T. S. Ingebrigtsen, T. B. Schroder, and J. Dyre. What is a simple liquid? *Physical Review X*, 2:011011, March 2012. 2

[18] T. D. Kühne and R. Z. Khaliullin. *Nature Communications*, 4:1450, 2013. 4

[19] T. D. Kuhne and R. Z. Khaliullin. Nature of the asymmetry in the hydrogen-bond networks of hexagonal ice and liquid water. *Journal of the American Chemical Society*, 136(9):3395–3399, 2014. 4

[20] T. Lazaridis and M. Karplus. *J. Chem. Phys.*, 105:4294, 1996. 3

[21] M. Leetmaa, K. T. Wikfeldt, M. P. Ljungberg, M. Odelius, J. Swenson, A. Nilsson, and L. G. Pettersson. Diffraction and ir/raman data do not prove tetrahedral water. *The Journal of chemical physics*, 129(8): 084502, 2008. 3

[22] M. Leocmach and H. Tanaka. Roles of icosahedral and crystal-like order in the hard spheres glass transition. *Nature communications*, 3: 974, 2012. 3

[23] Y. Liu and L. Wu. *J. Chem. Phys.*, 139:041103, 2013. 4

[24] R. L. McGreevy and L. Pusztai. Reverse monte carlo simulation: A new technique for the determination of disordered structures. *Mol. Simul.*, 1(6):359, 1988. 2

[25] A. Nilsson and L. G. M. Pettersson. *Chem. Phys.*, 389:1–34, 2011. 4

[26] L. QiKai and L. Mo. Rethinking atomic packing and cluster formation in metallic liquids and glasses. *Chinese Science Bulletin.*, 56(36):3897–3901, December 2011. 2

[27] M. Rovira Esteva. *Short-range order in disordered phases using neutron diffraction*. PhD thesis, Universitat Politècnica de Catalunya. Departament de Física Aplicada, 2014. 2

[28] M. Rovira-Esteva, L. C. Pardo, J. L. Tamarit, and F. J. Bermejo. Neutron diffraction as a tool to explore the free energy landscape in orientationally disordered phases. *Metastable Systems under Pressure*, pages 63–77, 2010. 3

[29] D. P. Shelton. Long-range orientation correlation in water. *The Journal of chemical physics*, 141(22):224506, 2014. 4

[30] A. Soper. Orientational correlation function for molecular liquids: the case of liquid water. *The Journal of chemical physics*, 101(8):6888–6901, 1994. 3

[31] A. Soper. An asymmetric model for water structure. *Journal of Physics: Condensed Matter*, 17(45):S3273, 2005. 3

[32] A. K. Soper. Empirical potential monte carlo simulation of fluid structure. *J. Chem. Phys.*, 202(2-3):295–306, January 1996. 2

[33] A. K. Soper. Partial structure factors from disordered materials diffraction data: An approach using empirical potential structure refinementempirical potential monte carlo simulation of fluid structure. *Phys Rev. B.*, 72(10):104204, September 2005. 2

[34] A. K. Soper. *Pure Appl. Chem.*, 82:1855, 2010. 4

[35] A. K. Soper and M. A. Ricci. Structures of high-density and low-density water. *Physical review letters*, 84(13):2881, 2000. 3

[36] F. Spaepen. Condensed-matter science: Five-fold symmetry in liquids. *Nature.*, 408:781–782, December 2000. 2

[37] F. H. Stillinger. *Science*, 209:451–457, 1980. 4

[38] I. Svishchev and P. Kusalik. Structure in liquid water: A study of spatial distribution functions. *The Journal of chemical physics*, 99(4): 3049–3058, 1993. 3

[39] Y. Takii, K. Koga, and H. Tanaka. A plastic phase of water from computer simulation. *The Journal of chemical physics*, 128(20):204501, 2008. 4

[40] H. Tanaka. Importance of many-body orientational correlations in the physical description of liquids. *Faraday discussions*, 167:9–76, 2013. 3

[41] K. T. Wikfeldt, M. Leetmaa, M. P. Ljungberg, A. Nilsson, and L. G. Pettersson. On the range of water structure models compatible with x-ray and neutron diffraction data. *The Journal of Physical Chemistry B*, 113(18):6246–6255, 2009. 3

[42] J. M. Ziman. *Models of disorder: the theoretical physics of homogeneously disordered systems*. CUP Archive, 1979. 2

Chapter 2

Computational methods

2.1 A brief overview of molecular dynamics

This section introduces a useful compilation of the physical and mathematical basis in which the molecular dynamics simulation method (MD) relies. For a detailed description of MD, I redirect the interested reader to the literature (3, 12).

2.1.1 The basis of the method

Molecular dynamics (MD) had its beginnings founded in the novel approach to solve a set of problems using statistical mechanics, introduced in a seminal paper by Metropolis et al. (20) in 1953. In 1957, Alder and Wainwright (1) tackled one open question, wether a hard sphere system could undergo on a solid-fluid phase transition using MD. This was an important work that set MD as a basic tool in statistical mechanics. Rahman used MD for the first time for a realistic fluid: a model of liquid Argon using Lennard-Jones particles in 1964 (27). Later in 1967 (36), Verlet introduced the now-famous integration Verlet algorithm, described in a later section. The MD method began as an application of the classical mechanics founded on Newton's Laws. So, it is classical in its heart, although nowadays there are several MD methods mixing quantum and classical mechanics after it was first introduced in 1976 by Warshel and Levitt (37). The scope of this thesis is limited to the use of the classical description of the atoms. On this basis

the rest of this chapter will focus on the classical method.

Atoms are the basic unit of MD. They are modeled as classical objects whose dynamic is ruled by the second law of movement of Isaac Newton, equation 2.1.

$$\mathbf{F}_i = m_i \frac{d^2 \boldsymbol{r}_i}{dt^2} = -\frac{\partial U(\boldsymbol{r})}{\partial \boldsymbol{r}_i} \tag{2.1}$$

The particle mass m_i is known. Then, the position vector $\boldsymbol{r}_i$ is obtained by integrating the acting forces upon the particle. The forces in MD are given by the gradient of the potential $U(\boldsymbol{r})$, this function contains all the physics for the model under study. Equation 2.2 shows the three terms that generally describe this function.

$$U(\boldsymbol{r}) = U_{bonded}(\boldsymbol{r}) + U_{non\text{-}bonded}(\boldsymbol{r}) + U_{special}(\boldsymbol{r}) \tag{2.2}$$

The first term considers the interactions within a molecule, such as the bond stretching, the angle bending, the torsional-angle rotation and the improper dihedral angle distortion that help to keep the geometry of some group of atoms.

$$\begin{aligned} U_{bonded}(\boldsymbol{r}) = & \sum_{bonds} \frac{1}{2} K_b (b - b_0)^2 + \sum_{\substack{bond \\ angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 \\ & + \sum_{\substack{improper \\ dihedral \\ angles}} \frac{1}{2} K_\xi (\xi - \xi_0)^2 + \sum_{\substack{dihedral \\ angles}} \frac{1}{2} K_\phi [1 + cos(m\phi - \delta)] \end{aligned} \tag{2.3}$$

In equation 2.3, letters b, θ, ξ and ϕ are the bond lengths, bond angles, improper dihedral angles and torsional dihedral angles respectively. These functions are harmonic, except for the torsional dihedral-angle term that has a trigonometric form. When the variables deviate away from their equilibrium values there is a penalty by increasing the potential energy so that those configurations are less probable to occur.

Non-bonded energies come from two terms, one that describes the van

der Waals interactions (35) with a Lennard-Jones function (15), and the other deals with the electrostatic interactions between partial charges of the atoms. The Lennard-Jones term accounts for the atomic repulsion from atom-atom overlap and the attraction from London dispersion forces (16).

$$U_{non\text{-}bonded}(\boldsymbol{r}) = \sum_{\substack{atom\\pairs}} \left[4\epsilon \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] + \frac{1}{4\pi\varepsilon_0} \sum_{\substack{atom\\pairs}} \frac{q_i q_j}{\varepsilon_1 r} \tag{2.4}$$

In equation 2.4 $\boldsymbol{r}$ is the distance between a pair of atoms, ϵ is the well depth of the potential and it governs the strength of the interaction. The value σ is the distance at which the potential is zero and it measures how close two particles can get. For the electrostatic term q_i and q_j are the particle charges, $4\pi\varepsilon_0$ is a constant containing the vacuum permittivity and ε_1 is the relative dielectric permittivity of the medium.

Finally, the special term $U_{special}(\boldsymbol{r})$ is optional. This term can contain an external potential, for example a restraining potential. For instance, a positional restraining can be used to obtain the initial orientationally disordered configurations for the plastic phases by keeping the positions and allowing the orientations to distribute randomly before the simulation starts.

The parameters included in the potential energy function are given by the so called *Force Fields*. A force field (FF) is a set of parameters given for every atom. These parameters are obtained by fitting to experimental data and there are a vast number of them, ranging from a wide range of biomolecules (17) to a set of specific molecules like water. It is important to note that FF's are designed to reproduce a set of properties and they can be valid only for sub-regions of the phase diagram. So, the FF must be chosen carefully according to the system and question to be solved.

2.1.2 The numerical tricks

2.1.2.1 The integration algorithm

Once the equation that governs the movement of the atoms is defined, the next step is to solve it. Equation 2.1 involves N particles and therefore it needs a numerical approach to be solved. Several integration algorithms are available nowadays, two of them are presented here: the Verlet algorithm (36) and the Leap-frog algorithm (see page 60 in (28)). The original Verlet algorithm can be derived using a Taylor expansion for the position $\boldsymbol{r}_i(t)$ of a particle forward in time:

$$\boldsymbol{r}_i(t+\Delta t) = \boldsymbol{r}_i(t) + \boldsymbol{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m}\Delta t^2 + \frac{\dddot{\boldsymbol{r}}}{3!}\Delta t^3 + O(\Delta t)^4 \tag{2.5}$$

And backward in time,

$$\boldsymbol{r}_i(t-\Delta t) = \boldsymbol{r}_i(t) - \boldsymbol{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m}\Delta t^2 - \frac{\dddot{\boldsymbol{r}}}{3!}\Delta t^3 + O(\Delta t)^4 \tag{2.6}$$

Adding these expressions gives

$$\boldsymbol{r}_i(t+\Delta t) \approx 2\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\Delta t) + \frac{\mathbf{F}_i(t)}{m}\Delta t^2 \tag{2.7}$$

It is interesting to note that equation 2.7 has a truncation error of fourth order $(\Delta t)^4$. This is because the odd terms are canceled out by the sum. Equation 2.7 can be used to obtain the time evolution of coordinates in MD with a small error. One drawback of the algorithm is that it does not define the velocities in an explicit manner. This can be a disadvantage when there is a need to know the kinetic energy of the system, for this another algorithmis needed, the Leap-frog method.

The Leap-frog algorithm can be derived in a similar way to Verlet's. Equation 2.5 can be rewritten as:

$$\boldsymbol{r}_i(t+\Delta t) = \boldsymbol{r}_i(t) + \Delta t\left[\boldsymbol{v}_i(t) + \frac{\Delta t}{2}\frac{\mathbf{F}_i(t)}{m}\right] + O(\Delta t)^3 \tag{2.8}$$

Where the terms within the bracket are the Taylor expansion of the velocity using a half-time step. This leads to the expression,

$$\boldsymbol{r}_i(t+\Delta t) = \boldsymbol{r}_i(t) + \boldsymbol{v}_i\left(t+\frac{\Delta t}{2}\right)\Delta t \tag{2.9}$$

But now, the velocity at half-time step is needed. This can be derived using a Taylor expansion for the velocity $\boldsymbol{v}_i(t)$ of a particle forward in time:

$$\boldsymbol{v}_i\left(t+\frac{\Delta t}{2}\right) = \boldsymbol{v}_i(t) + \frac{\mathbf{F}_i(t)}{m}\frac{\Delta t}{2} + \frac{\ddot{\boldsymbol{v}}}{2}\left(\frac{\Delta t}{2}\right)^2 + O\left(\frac{\Delta t}{2}\right)^3 \tag{2.10}$$

and backward in time,

$$\boldsymbol{v}_i\left(t-\frac{\Delta t}{2}\right) = \boldsymbol{v}_i(t) - \frac{\mathbf{F}_i(t)}{m}\frac{\Delta t}{2} + \frac{\ddot{\boldsymbol{v}}}{2}\left(\frac{\Delta t}{2}\right)^2 - O\left(\frac{\Delta t}{2}\right)^3 \tag{2.11}$$

Subtracting the equations gives,

$$\boldsymbol{v}_i\left(t+\frac{\Delta t}{2}\right) = \boldsymbol{v}_i\left(t-\frac{\Delta t}{2}\right) + \frac{\mathbf{F}_i(t)}{m}\Delta t + O\left(\frac{\Delta t}{2}\right)^3 \tag{2.12}$$

Equations 2.9 and 2.12 constitute the Leap-frog algorithm. Positions and velocities are computed at different times, from there the name of the method. One advantage of the method is that the velocity is now available. One drawback it has is that now the truncation error is of order three, lower than Verlet's. Truncation errors are also known as *local errors*, which are caused by one iteration. But the cumulative errors are even more important, these are called *global errors*. The Verlet and Leap-frog methods have a second order global error. The Leap-frog integration method was used in this thesis.

The election of the time step Δt is crucial in MD to avoid important accumulation of errors. The value of Δt should be less than the characteristic time of the molecular vibrations under study, but it also has to be big enough for recording a large set of microstates. This allows the properties studied to be statistically significant. A value of 1 fs is a typical value.

And for small molecules systems the simulations are of the order of 1 ns. This means producing 10^6 microstates to allow the simulation to represent statistically the real system.

2.1.2.2 Periodic boundary conditions

One of the limitations in MD simulations is the small number of particles considered to solve the equations compared to a real system. Alder & Wainwright performed the first MD simulation for a system of 32 hard spheres (2). By the time of writing, N $\approx 10^3$ is a typical system size while the largest simulation consisted of N $\approx 10^{12}$ particles (7). Although impressive, that simulation is yet far away from real systems with the Avogadro's constant N $\approx 10^{23}$ number of particles. For instance, a simulation of a simple cubic crystal with 1000 particles will have 512 of them at the surface. This creates a problem to compute the bulk properties. Periodic boundary conditions (PBC's) are used to overcome this finite size problem.

In a simulation with PBC's the atoms can interact with their direct neighbours and also with the periodic images of those atoms, as shown in figure 2.1. The assumption made by using the PBC's is that the physical boundaries are far away, enough to not influence the bulk behaviour. This is valid for most macroscopically large systems and also for the systems of interest of this thesis.

Since PBC's are not a real description of the system, some considerations have to be taken into account: atoms that leave the box re-enter through the opposite face as it is indicated by arrows in figure 2.1 and the atoms should not interact with their replica images. This is known as the *minimum image convention*. This last requirement has a consequence to the use of *cut-off's* in MD. This means that non-bonded interactions are computed only within a predefined distance, which should be less than half of the box distance to respect the minimum image convention.

The use of *cut-off's* play another important role: they save computing time. It would be costly to compute all the interactions between parti-

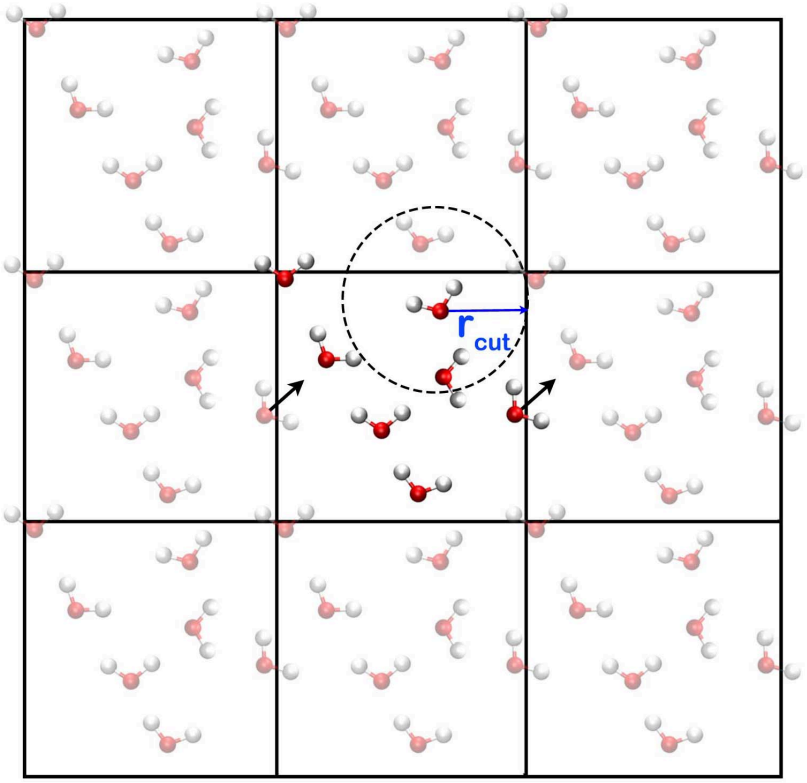

Figure 2.1: Periodic boundary conditions used in MD simulations. The distance r_{cut} shows the limit of the interaction introduced by the *cut-off*. The simulation box corresponds to a liquid water example.

cles. Lennard-Jones(LJ) interactions decay with distance as $1/\boldsymbol{r}^6$, figure 2.2 shows its functional form. The LJ potential is usually truncated at a distance of 2.5σ, where it is close to zero. There are several truncation forms, two of them are shown in figure 2.2. The differences between them arise only by zooming into the cut-off region. The first one is a plain cut-off, for which the LJ potential goes to zero discontinuously. The second one is known as a Switch cut-off, the LJ decays to zero beyond the cut-off distance in a continuous way. The later has the advantage that it does not introduce discontinuities in the energy.

The electrostatic interactions have a different treatment. The decay of this function is slower than van der Waals as it goes as $1/\boldsymbol{r}$. For this

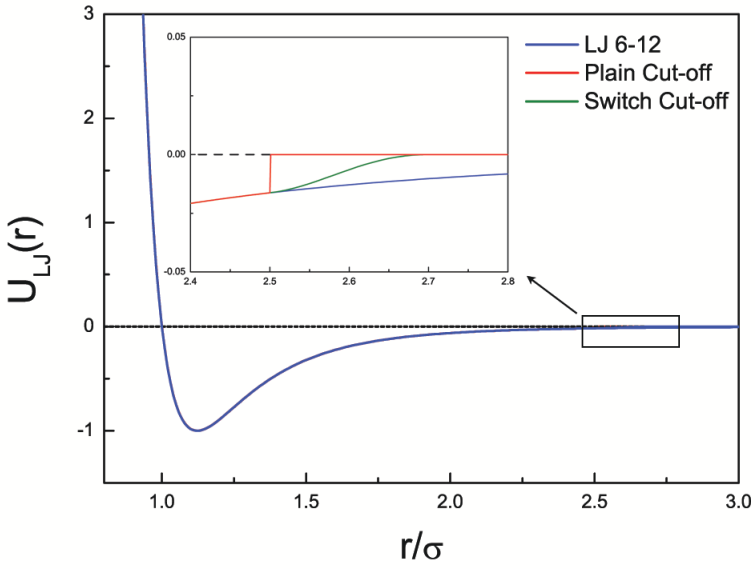

Figure 2.2: Cut-off for the Lennard-Jones interactions. A zoom around the typical cut-off value of 2.5σ is shown for two schemes: a plain cut-off and a Switched cut-off.

kind of long-range interactions introducing a cut-off would produce serious errors. It is therefore necessary to account for all the interactions within the simulation box and its replicas. One of the methods to solve this problem is studied in the next section.

2.1.2.3 The Ewald summation for long range potentials

The Ewald summation is a method designed for evaluating the potential of an N particles system under periodic boundary conditions (8),

$$U = \frac{1}{2} \sum_{i,j=1}^{N} \sum_{\boldsymbol{n}}{}' u(\boldsymbol{r}_{ij} + \boldsymbol{n}L) \tag{2.13}$$

In equation 2.13, the sum over $\boldsymbol{n}$ counts all the periodic replicas of the particles, the prime indicates that for $\boldsymbol{n} = 0$ the case i = j must be neglected, and $\boldsymbol{r}_{ij}$ is the distance vector between particles. This method is used in MD to evaluate the Coulomb potential according to equation 2.14:

$$U_{coulomb} = \frac{1}{2} \sum_{i,j=1}^{N} \sum_{\boldsymbol{n}}^{'} \frac{q_i q_j}{\boldsymbol{r}_{ij} + \boldsymbol{n}L} \tag{2.14}$$

This sum is conditionally convergent (34), physically due to the slowly decaying long-range of the Coulombic interactions. So, a direct evaluation of this sum is problematic. The Ewald summation method split the sum in equation 2.14 into a sum of two rapidly converging series in real and reciprocal space of the form,

$$\sum \frac{\boldsymbol{F}(n)}{\boldsymbol{n}} + \sum \frac{1 - \boldsymbol{F}(m)}{\boldsymbol{m}} \tag{2.15}$$

The function $\boldsymbol{F}$ is chosen so that as $\boldsymbol{n} \to \infty$, $\boldsymbol{F}(n)$ decays rapidly and the series converges also rapidly. Besides, the second part will be a slowly varying function (i.e. a smooth function) and hence its Fourier transform decays rapidly. The Ewald sum is therefore written as the sum of those series plus a constant term,

$$U_{Ewald} = U^{(r)} + U^{(k)} + U^{(s)} \tag{2.16}$$

The mathematical form of the contributions from real space $U^{(r)}$, reciprocal space $U^{(k)}$ and a self-energy correction term $U^{(s)}$ are described in the article from Toukmaji and Board Jr. (34). The next paragraph explains their physical meaning.

Each point charge in the simulation system can be seen as a delta function. Each charge can be surrounded by a Gaussian charge distribution of equal magnitude and opposite sign as it is shown in figure 2.3. This effectively limits the interactions to a short range. As this function is a new introduced term, it has to be counteracted. In order to do this, a second Gaussian charge distribution is introduced with the same sign and magnitude as the original point charge. The sum is now performed in the reciprocal space using Fourier transforms. Finally, the self-energy terms add a correction term. It is also important to note that the function in the Ewald method is usually chosen different to a Gaussian function.

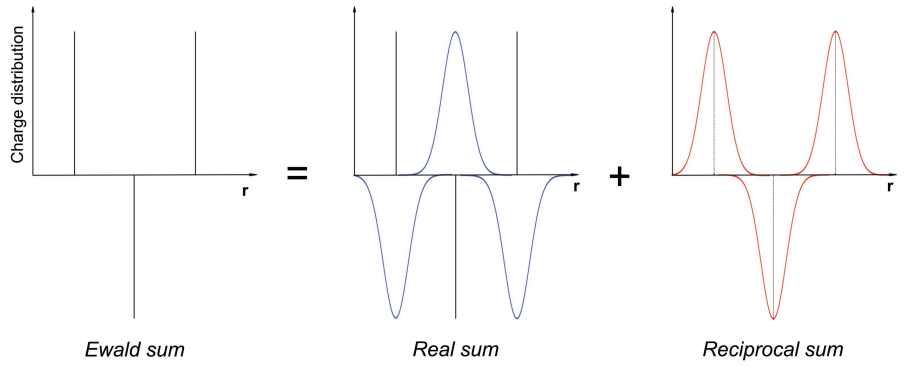

Figure 2.3: The Ewald summation method. The sum is divided in the real and reciprocal spaces with the aid of an added function, a gaussian in this example that screens the point charges.

The Ewald summation method has one drawback: its computational cost. It requires a direct evaluation that scales with the number of particles as N^2. There exist a number of methods that decompose the sum on a grid (or 'mesh'). The improvement lays on the evaluation of the reciprocal sum using a fast Fourier transform. The Particle Mesh Ewald Method (6) was used in this thesis. It scales as $N{\cdot}log(N)$.

2.1.3 A typical MD simulation

This section reviews the basic steps to perform an MD simulation. The results obtained in this thesis are concerned with the equilibrium properties of the system. This means that the system is evolved from an initial state (the positions and velocities of the N particles). The first step is then,

1. Build the initial configuration or positions of the system

The next step consists in assigning velocities, as they are needed for the Leap-frog integrator, see equation 2.12.

2. Initial velocities are assigned according to a Maxwell-Boltzmann distribution at a desired temperature.

The initial state can be evolved in time and the forces updated every time step of integration.

3. Compute forces acting on particles to update positions and velocities.

In MD, like in real experiments some parameters can be controlled, like temperature or pressure. This is done via the so called *thermostats* and *barostats*. A *thermostat* controls the temperature by coupling the system to an external reservoir. Numerically, this is done basically by rescaling the velocities every n-steps of simulation. Increasing the velocities when the system is under the desired temperature T_0 or decreasing them when the system is above T_0. Equation 2.17 shows the Berendsen thermostat (4)

$$\frac{dT}{dt} = \frac{1}{\tau_T}(T_0 - T) \tag{2.17}$$

The coupling value τ_T indicates the frequency of velocity rescaling. It is typically of the order of 1 ps.

In a similar way a *barostat* controls the pressure of the system. In this case, the box-size, particle positions and momenta are rescaled. In a simple way, the box length is increased and therefore the volume of the box increases lowering the pressure. The simplest rescaling is isotropic, although there are also anisotropic methods. Equation 2.18 shows the Berendsen barostat (4)

$$\frac{dP}{dt} = \frac{1}{\tau_P}(P_0 - P) \tag{2.18}$$

The Berendsen coupling is a good method to equilibrate the system at the desired temperature and pressure, although it does not generate a correct thermodynamic ensemble. The Nose-Hoover thermostat (22) and Parrinello-Rahman barostat (24) solve this issue. A common practice is to begin the simulation with the Berendsen coupling and switch to the later thermostat and barostat for the production run to record the configurations for further analyses.

4. Rescaling on positions and velocities according to thermal and pressure coupling.

Finally, the configurations are recorded and the properties of interest are evaluated.

5. Record the new configuration state and compute thermodynamical properties.

There are several packages for performing MD simulations, either under commercial or free license. The Gromacs engine (GROningen MAChine Simulation) (14) was used for all the simulations performed in this thesis.

2.1.4 Observable from MD

In a real experiment the measurements are done for the macroscopical properties of the system and not for the positions and velocities. The real system can pass through a huge number of conformations. In statistical mechanics, the properties measured correspond to ensemble averages. To which ensemble corresponds most closely the experiment is irrelevant given that there is a large enough number of particles to be close of the thermodynamic limit ($N \approx 10^{23} \rightarrow \infty$).

In an MD simulation the system passes through microscopic configurations at each time step. So, the points in the ensemble are computed sequentially in time and the system must pass through all the possible states to calculate an ensemble average. But it is necessary to have a connection between the ensemble average of a real system and the time average produced in MD. The *Ergodic Hypothesis* defines this connection and is one of the most fundamental axioms of statistical mechanics.

$$\langle a \rangle_{ensemble} = \langle a \rangle_{t \rightarrow \infty} \tag{2.19}$$

Equation 2.19 shows that the ensemble average for a given property a can be evaluated in time if one allows the system to evolve in time indefinitely such that it passes through all possible states. In MD this is not strictly the case given that the simulation time is limited by computer resources. It is nevertheless a good approximation and one must be certain to sample a sufficient amount of phase space according to the system under study.

Throughout this thesis different observable derived from the positions

and velocities of the particles were studied. Mainly, those related to the structure or static properties. And also in chapter 7 some dynamical properties for different structural phases of water were computed.

2.2 ANGULA: An analysis tool into the ordering of molecules of disordered phases

The MD method offers an insight into the atomic details that is difficult or impossible for experimental techniques to reach. The basic output of an MD simulation is the set of positions and velocities of particles. The analysis of that data can lead to comparison with experiments or to theoretical insights. This section presents ANGULA: an analysis tool that uses the Euler angles to characterize the way in which the molecules are ordered.

2.2.1 The idea behind ANGULA

Contrary to what happens in an ordered solid crystalline phase where it is necessary only to study the unit cell to reproduce the long (infinity) range structure, in liquids it is necessary to study the whole system to get some information about the most probable arranging of molecules (position and orientation). Considering molecules as rigid bodies in three dimensional space, they have six degrees of freedom: three positional and three rotational. By attaching an axis set to every molecule in the system the arranging of any molecule relative to another reference molecule can be fully described. Finally, the distribution of probable positions and orientations is obtained to characterize the structure of the system. This method is valuable to extract information from disordered phases, such as the liquid and plastic phases that are the subject of this thesis.

Lets detail the way in which ANGULA works: An axis set is attached to a molecule, the relative position of a second molecule with the same axis set is known from the distance (r) between the origins of the axis sets and two angles (θ_{pos} and ϕ_{pos}) as it is shown in figure 2.4. The relative orientation of the second rigid body to the first one can be known by using three Euler angles (θ_{ori}, ϕ_{ori} and ψ_{ori}). The convention ZY′Z″ is used to obtain the

Euler angles, which means that the operations needed to get the rotated axis of the second molecule from the original one are: a rotation around the Z-axis (ϕ_{ori}), followed by a rotation around the new Y′-axis (θ_{ori}) and finally a rotation around the new Z″-axis (ψ_{ori}).

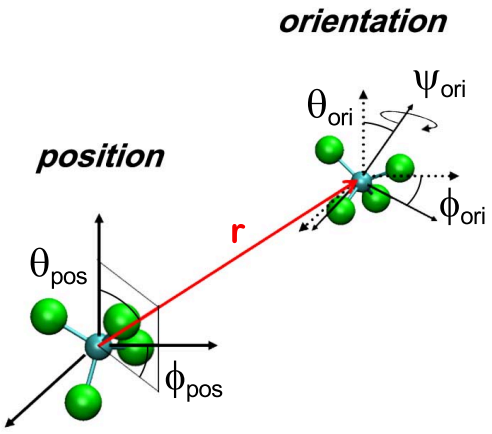

Figure 2.4: Definition of the Euler angles used to characterize the position and orientation of two molecules.

For a system with n-molecules, the distribution of probable arrangements for the first neighbour of a molecule can be obtained. Another example is to obtain the distribution for the arrangement around a certain distance (r). In the first case, the nearest molecule is analyzed and the results are averaged over every molecules and through all the possible frames of simulation. In the latter case, all the molecules within the distance (r) are taken into account. The set of probable arrangements is defined by $g(\Omega_{\text{pos}}, \Omega_{\text{ori}}, \text{r})$, being Ω_{pos} the set of possible angles θ_{pos} and ϕ_{pos}, that are the angles that define the position of a molecule in spherical coordinates. While Ω_{ori} is the set of possible Euler angles θ_{ori}, ϕ_{ori} and ψ_{ori}.

2.2.2 A working example

Some of the results that can be obtained using ANGULA are presented here. To do that, the liquid carbon tetrachloride is chosen as an example as it was the first molecular liquid to be studied by means of X-ray diffraction

technique and its relative molecular ordering has been studied before (21, 29), defining a tetrahedral symmetry.

2.2.2.1 MD simulation

In order to obtain the molecular configurations necessary to calculate $g(\Omega_{pos}, \Omega_{or}, r)$ an MD simulation using the Gromacs 4.5 (13) package was performed. The potential parameters used were for carbon $\sigma = 3.7746$Å, $\epsilon = 0.2271$kJ/mol, q $= -0.696$e and for chloride: $\sigma = 3.4667$Å, $\epsilon = 1.0944$kJ/mol and q $= 0.174$e. The simulation was made on a 216 rigid molecules system using the NPT ensemble at the thermodynamics conditions of the liquid, namely T = 298K and P = 1atm.

The chosen time step to perform the simulation was Δt = 5fs (during 1200ps), switched cut-offs from 8Å to 14Å for Lennard-Jones interactions and 14Å for coulomb pairs were used. The Particle Mesh Ewald (PME) method was applied beyond the electrostatic cut-off in order to obtain the reciprocal space sum.

2.2.2.2 Axes definition

The first step consists of defining the axes definition for the CCl_4 molecule. Technically, this is done by adding four "pseudo-atoms" below the coordinates of each molecule in a position file (PDB, XYZ or any other). The "pseudo-atoms" are: the origin of the axis set, X, Y and Z. The axes chosen are shown in figure 2.4: the Z axis is set in the bisecting angle of two C-Cl bonds, the Y axis is set perpendicular to this axis and coplanar with a Cl-C-Cl plane, and the X axis is determined as usually in an orthonormal axis set ($X = Y \times Z$). The origin of the system is located in the Carbon atom.

2.2.2.3 Short range order analysis

Figure 2.5 shows an example of the output file. Reading The fifth line: the first two columns show the molecules involved in the calculation (the 9th and 173th), corresponding to the central and neighbour molecules. Next,

```
Angle file from: 1000 configurations
         9          10
   central   neighbour  d           cos(theori)   phiori       psiori    cos(thepos)   phipos
   *******   Conf       =           1
      9        173      4.825808    -0.31255     124.8514    -161.1203   -0.8460      -113.5446
     10        110      5.039682     0.16398    -141.0552      86.0169   -0.8809       -94.43398
       .
       .
       .
```

Figure 2.5: Output file from an analysis of ANGULA.

the distance d between the two axis set is shown. The Euler angles defining the position are thepos and phipos, while theori, phiori and psiori are related with the relative orientation.

The next step is to calculate the positional part of $g\left(r, \Omega_{pos}, \Omega_{ori}\right)$. For this example the distance r corresponding with the first four neighbours was chosen in order to characterize the tetrahedral ordering around a molecule. The panel a1) of figure 2.6 shows the two-dimensional probability distribution functions (PDF) $g\left(\cos(\theta_{pos}), \phi_{pos}\right)$ describing the molecular position. Panel a2) of the same figure shows the spatial density map of the high probability regions of $g\left(\cos(\theta_{pos}), \phi_{pos}\right)$. The tetrahedral ordering can be seen in the spatial density map: neighbouring molecules are located in the faces of the pyramid formed by the chloride atoms.

In order to study the molecular orientation the highest probability regions of $g\left(\cos(\theta_{pos}), \phi_{pos}\right)$ were chosen, i.e. some specific molecular positions. Panels b1) and b2) of figure 2.6 show the 3D probability distribution for the Euler angles describing the molecular orientation for molecules in the highest probability positions labeled as b1) and b2) in $g\left(\cos(\theta_{pos}), \phi_{pos}\right)$. Without entering into details (see reference (26)), the obtained spirals represent a series of probable orientations that correspond to a molecule rotating around its C-Cl axis with its face parallel to the central molecule face. Panel b3) of figure 2.6 shows the complete 3D probability distribution. As it can

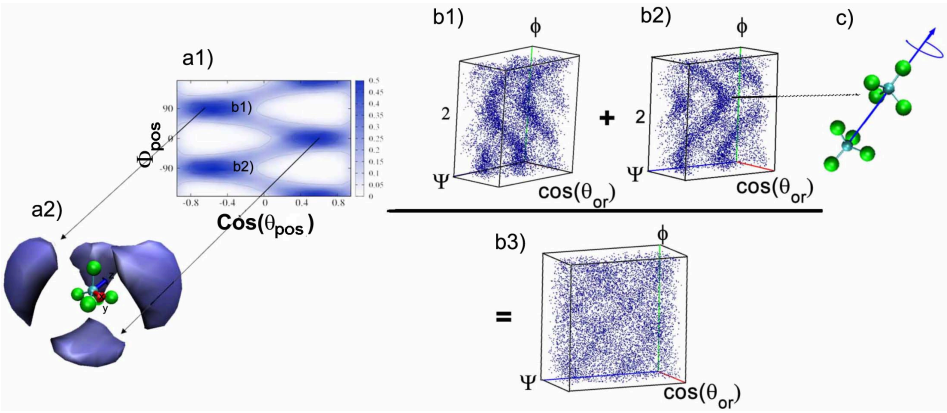

Figure 2.6: Panel a1) shows the probability distribution of the angles that determine molecular position $\cos(\theta_{CM})$ and ϕ_{CM} for the closest four molecules to a central one. Panel a2) shows the spatial density map obtained from the probability map a1). Panels b1) and b2) show the 3D orientational probability distributions of the angles determining molecular orientation θ_{or}, ϕ_{or} and ψ_{or} for different regions of the positional map shown in panel a1). Panel b3) shows the merging of all orientational probability distributions $g(\Omega_{or})$.

be readily seen, it is difficult to extract any detailed features concerning molecular orientations from this distribution function. At least, it is a time consuming task for the user to analyze every possible ordering. In order to get some general information from these 3D maps the concepts of information theory will be helpful, this is shown in chapter 3.

2.3 Bayesian fit using Markov Chain Monte Carlo

This section describes a methodology to obtain the intra molecular structure from experimental data. The intra molecular structure is a necessary input for MD or refinement methods such as reverse Monte Carlo (19) and empirical structure refinement (33). This methodology is grounded in Bayesian theory, performing a fit to the experimental data that in our case is a neutron diffraction experiment that contains information on the structure of molecular systems.

The methodology of this bayesian fit is implemented in the code FABADA (Fitting Algorithm for Bayesian Analysis of DAta) (23). FABADA is a mul-

tipurpose fitting algorithm, it has been used for diffraction (30), dielectric spectroscopy (18), astrophysics (31) and quasielectric neutron scattering (5). FABADA was used in this thesis to fit the intramolecular structure of hexachloroethane obtained from neutron diffraction experiments. The concepts in which FABADA relies are presented next.

2.3.1 Important concepts

FABADA works in the same way as a Monte Carlo simulation, using the fitting merit function as an equivalent of the energy: $\mathrm{E} = \chi^2$. It generates successive parameter sets from a model H (also called hypothesis) constructed to describe the experimental data D. The probability that a parameter set from the model H is in fact describing the data, i.e. the conditional probability $\mathrm{P}(D \mid H)$, can be calculated assuming that experimental data points D_k are normally distributed around their expected value H_k with a standard deviation σ_k:

$$\begin{aligned} \mathrm{P}\left(D \mid H\right) &\propto \prod_{k=1}^{n} \exp\left[-\frac{1}{2}\left(\frac{H_k - D_k}{\sigma_k}\right)^2\right] \\ &= \exp\left[-\frac{1}{2}\sum_{k=1}^{n}\left(\frac{H_k - D_k}{\sigma_k}\right)^2\right] \\ &= \exp\left(-\frac{\chi^2}{2}\right) \end{aligned} \tag{2.20}$$

where k designates each particular point, n is the total number of measured points, and σ_k is the experimental error associated with each k.

A new parameter set that improves the fit will always be accepted, but if it worsens the fit the set will only be accepted with a probability:

$$\frac{\mathbb{P}\left(H\{P_i^{\mathrm{l}+1}\} \mid D\right)}{\mathbb{P}\left(H\{P_i^{\mathrm{l}}\} \mid D\right)} = \exp\left(-\frac{\chi^2_{\mathrm{l}+1} - \chi^2_{\mathrm{l}}}{2}\right) \tag{2.21}$$

where $\mathrm{P}(H \mid D)$ is the conditional probability that the hypothesis H is

true given the experimental data D, $H\{P_i^{\mathrm{l}}\}$ is the model with the old set of parameters and $H\{P_i^{\mathrm{l}+1}\}$ with the new set of parameters, D_k are the experimental points, and χ^2_{l} and $\chi^2_{\mathrm{l}+1}$ are the χ^2 associated to the old and new set of parameters, respectively.

Another interesting feature is that "previous knowledge" of the system can be used to help with the fitting, this is known as a prior. Prior information I can be included by a redefinition of χ^2 as proposed in (32). Following that work, suppose that some parameters were previously measured obtaining the values $X_j = x_{oj} \pm \epsilon_j$, where j runs for the number (M) of parameters and ϵ accounts for their errors. If an uncorrelated Gaussian PDF for the prior for X is assumed then,

$$prob(\mathbf{X}|I) = \prod_j^M \frac{1}{\epsilon_j\sqrt{2\pi}} exp\left[-\frac{(X_j - x_{oj})^2}{2\epsilon_j^2}\right] \propto exp\left(-\frac{C}{2}\right) \tag{2.22}$$

where

$$C = \sum_j^M \left(\frac{X_j - x_{oj}}{\epsilon_j}\right)^2 \tag{2.23}$$

therefore, the addition of the factor C to the definition of χ^2 allows to account for previous experimental information of the parameters.

2.3.2 Fitting the structure factor and the radial distribution function

A structure factor corresponding to the contribution of a pair of atoms from the molecule being separated a characteristic distance d, and having a Debye-Waller term $l^2 \equiv \left\langle \delta r_{ij}^2 \right\rangle = \langle u_i^2 \rangle + \left\langle u_j^2 \right\rangle$, where $\langle u_k^2 \rangle$ $(k = i, j)$ is the mean squared vibrational amplitude for the atom at site k, can be approximated as (11):

$$S_{test}(q) = \frac{\sin(qd)}{(qd)} \exp\left[-\frac{1}{2}(lq)^2\right] \tag{2.24}$$

The result of a Bayesian fit is a collection of all parameter PDFs, which can be simply obtained by counting the recurrence of the parameter combinations that were accepted during the MCMC procedure.

The Nyquist-Shannon (NS) sampling theorem applied to diffraction experiments establishes an inferior limit on the spacing of the real space function:

$$\delta r = \frac{\pi}{q_{max}} \tag{2.25}$$

where δr is the interval with which the real space is sampled. In the article of Farrow et al. (9) a nice explanation of this theorem and its implications avoiding oversampled pair distribution functions can be found. Summarizing their conclusions, although an oversampled PDF "looks nicer" it has redundant information. In other words, fits can be performed with as few points as those given by the NS theorem without any loss of information.

In real diffraction instruments, due to the fact that the experimental q range (q_{exp}) is limited, before comparing a modelled radial distribution function (RDF) with the Fourier transform of the experimental structure factor, the modelled RDF must be convoluted with the Fourier transform of the window function, to properly account for its contribution to the experimental RDF. However, convolution of functions is a time consuming process for any computing algorithm. An interesting property of NS sampling is that when the RDF is not oversampled, no convolution is required to fit the Fourier transformed experimental data in real space: The spurious broadening of the peaks due to the limited q-range is automatically taken into account by a coarsening of the points in the RDF, and the contribution of the ripples appearing at the sides of a peak is also naturally eliminated. This is due to the functional form of the Fourier transform of the window function that must be used in the real space (11):

$$M(r) = \frac{1}{\pi}\left[\frac{\sin\left(q_{\max} r\right)}{r}\right] \tag{2.26}$$

Using a NS sampling such as $r_i = i\,\delta r$ (where $i = 0, \ldots, n$) that complies with the interval $\delta r = \pi / q_{max}$ mentioned before, and computing the con-

volution function $M(r)$, it becomes clear that, except for the first point, $r_0 = 0$, the convolution function is always crossing zero at the r_i points when using this NS interval:

$$M(r_i) = \frac{1}{\pi}\left[\frac{\sin(q_{\max} i\,\delta r)}{i\,\delta r}\right] = \frac{q_{max}}{\pi}\left[\frac{\sin(i\,\pi)}{i\,\pi}\right] = \begin{cases} q_{max}/\pi & for\, i = 0 \\ 0 & for\, i \neq 0 \end{cases} \tag{2.27}$$

Therefore turning the laborious convolution process into a superfluous task that is no longer necessary. As it has been stated before, fitting both the real and reciprocal spaces can be advantageous, therefore without losing data from the experiment, one can make the fitting faster and more reliable using NS theorem by not oversampling the real space. The use of a bayesian fitting is shown in section 4.3.

2.4 Convolution of the MD $S(q)$ with the experimental error function

The static structure factor $S(q)$ obtained from MD needs to be corrected to account for the experimental broadening of the peaks due to the limitations in the resolution of the instrument. The details of this calculation are shown in this section, which will be used in chapter 4.

Figure 2.7 shows the q-resolution Δq as a function of q obtained for a calibration sample, a quadratic fit to this function was done and it is shown in equation 2.28. Taking the instrumental resolution function $I(q)$ as a gaussian function with a FWHM determined by the q-resolution function (25) (see the inset in Figure 2.7 for two values of q).

$$\Delta(q) = 0.00584q^2 - 0.04057q + 0.18696 \tag{2.28}$$

Next, the convolution of the $S(q)$ with $I(q)$ was computed, i.e, a gaussian function centered at a given value q with standard deviation given by the relationship $FWHM \approx 2.3548\ \sigma$, obtaining finally the static structure factor comparable to the experiment $S_{out}(q)$:

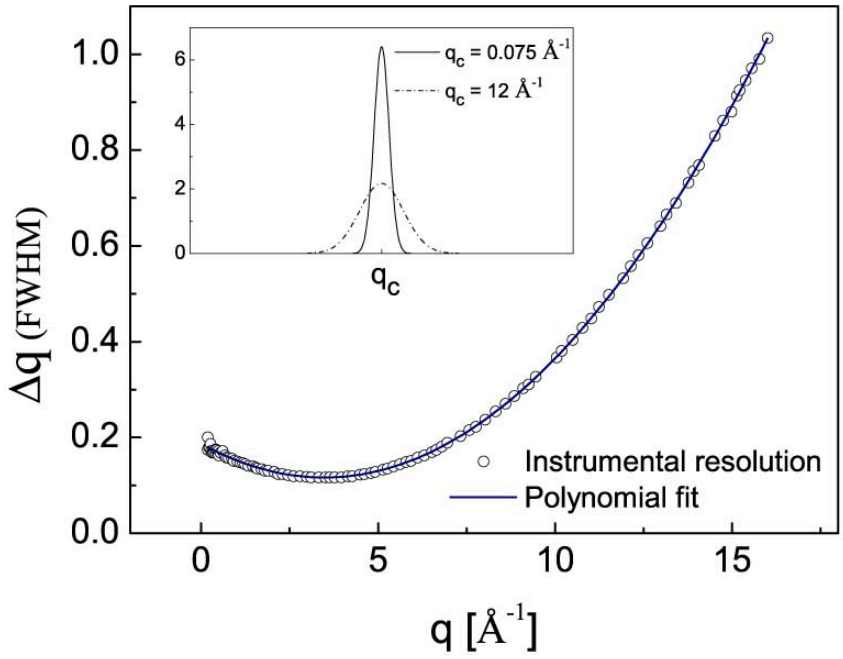

Figure 2.7: Instrumental resolution function Δq as a function of q from the D4c3 (10) diffractometer. Dots are the experimental values from the calibration sample and the thick line shows the polynomial fit (equation 2.28). The inset shows two gaussian functions with a FWHM determined by the value of Δq at a given q (q_c). At higher q values, the gaussian function becomes wider.

$$S_{out}(q) = \sum_{n}^{Np} S(n)I(n-q) \tag{2.29}$$

$$S_{out}(q) = \sum_{n}^{Np} S(n)\frac{2.3548}{2FWHM}\exp^{-\frac{1}{2}\left(\frac{2.3548(n-q)}{FWHM}\right)^2} \tag{2.30}$$

where $S(n)$ is the original MD static structure factor and n runs for every point in $I(n)$, in this case from $0.075\mathring{A}^{-1}$ to $14\mathring{A}^{-1}$ every $0.025\mathring{A}^{-1}$, and the value of $FWHM$ is given by $\Delta(q)$, so it can be rewritten:

$$S_{out}(q) = \sum_{n}^{Np} S(n) \frac{2.3548}{2\Delta(q)} \exp^{-\frac{1}{2}\left(\frac{2.3548(n-q)}{\Delta(q)}\right)^2} \tag{2.31}$$

2.5 Bibliography

[1] B. Alder and T. Wainwright. Phase transition for a hard sphere system. *The Journal of chemical physics*, 27(5):1208, 1957. 11

[2] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. i. general method. *J. Chem. Phys.*, 31:459–466, 1959. 16

[3] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, New York, first edition, 1987. 11

[4] H. J. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81(8):3684–3690, 1984. 21

[5] S. Busch, C. Smuda, L. C. Pardo, and T. Unruh. Molecular mechanism of long-range diffusion in phospholipid membranes studied by quasielastic neutron scattering. *Journal of the American Chemical Society*, 132 (10):3232–3233, 2010. 28

[6] T. Darden, D. York, and L. Pedersen. Particle mesh ewald: An $N{\cdot}log(N)$ method for ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092, 1993. 20

[7] W. Eckhardt, A. Heinecke, R. Bader, M. Brehm, N. Hammer, H. Huber, H.-G. Kleinhenz, J. Vrabec, H. Hasse, M. Horsch, et al. 591 tflops multi-trillion particles simulation on supermuc. In *Supercomputing*, pages 1–12. Springer, 2013. 16

[8] P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Ann. Phys.*, 369:253–287, 1921. 18

[9] C. L. Farrow, M. Shaw, H. Kim, P. Juhás, and S. J. Billinge. Nyquist-shannon sampling theorem applied to refinements of the atomic pair distribution function. *Physical Review B*, 84(13):134105, 2011. 30

[10] H. E. Fischer, G. J. Cuello, P. Palleau, D. Feltin, A. C. Barnes, Y. S. Badyal, and J. M. Simonson. *Appl. Phys. A: Mater. Sci. Process.*, 74: S160, 2002. 32

[11] H. E. Fischer, A. C. Barnes, and P. S. Salmon. Neutron and x-ray diffraction studies of liquids and glasses. *Reports on Progress in Physics*, 69(1):233, 2005. 29, 30

[12] D. Frenkel and B. Smith. *Understanding Molecular Simulation*. Oxford University Press, London, second edition, 2001. 11

[13] B. Hess, C. Kutzner, D. V. der Spoel, and E. Lindahl. Gromacs4: Algorithms for highly efficient, load-balanced and scalable molecular simulation. *J. Chem. Theory Comput.*, 4:435–447, 2008. 25

[14] B. Hess, C. Kutzner, D. Van Der Spoel, and E. Lindahl. Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3):435–447, 2008. 22

[15] J. E. Lennard-Jones. On the determination of molecular fields. ii. from the equation of state of a gas. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 106, pages 463–477. The Royal Society, 1924. 13

[16] F. London. Über einige eigenschaften und anwendungen der molekularkräfte. *Z. phys. Chem*, 11:222–251, 1930. 13

[17] A. D. Mackerell. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.*, 25(13):1584–1604, 2004. 13

[18] J. C. Martinez-Garcia, J. L. Tamarit, L. C. Pardo, M. Barrio, S. J. Rzoska, and A. Droz-Rzoska. Disentangling the secondary relaxations in the orientationally disordered mixed crystals: cycloheptanol+ cyclooctanol two-component system. *The Journal of Physical Chemistry B*, 114(18):6099–6106, 2010. 28

[19] R. L. McGreevy and L. Pusztai. Reverse monte carlo simulation: A new technique for the determination of disordered structures. *Mol. Simul.*, 1(6):359, 1988. 27

[20] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953. 11

[21] A. Narten, M. Danford, and H. Levy. Structure and intermolecular potential of liquid carbon tetrachloride derived from x-ray diffraction data. *The Journal of Chemical Physics*, 46(12):4875–4880, 1967. 25

[22] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984. 21

[23] L. Pardo, M. Rovira-Esteva, S. Busch, J.-F. Moulin, and J. L. Tamarit. Fitting in a complex χ 2 landscape using an optimized hypersurface sampling. *Physical Review E*, 84(4):046711, 2011. 27

[24] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52(12):7182–7190, 1981. 21

[25] J. S. Pedersen, D. Posselt, and K. Mortensen. Analytical treatment of the resolution function for small-angle scattering. *J. Appl. Cryst.*, 23: 321–333, 1990. 31

[26] S. Pothoczki, L. Temleitner, P. Jóvári, S. Kohara, and L. Pusztai. Nanometer range correlations between molecular orientations in liquids of molecules with perfect tetrahedral shape: Ccl4, sicl4, gecl4, and sncl4. *The Journal of chemical physics*, 130(6):064503, 2009. 26

[27] A. Rahman. Correlations in the motion of atoms in liquid argon. *Physical Review*, 136(2A):A405, 1964. 11

[28] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, Cambridge, second edition, 2004. 14

[29] R. Rey. Quantitative characterization of orientational order in liquid carbon tetrachloride. *The Journal of chemical physics*, 126(16):164506, 2007. 25

[30] M. Rovira-Esteva, N. Murugan, L. Pardo, S. Busch, J. L. Tamarit, S. Pothoczki, G. Cuello, and F. J. Bermejo. Interplay between intramolecular and intermolecular structures of 1, 1, 2, 2-tetrachloro-1, 2-difluoroethane. *Physical Review B*, 84(6):064202, 2011. 28

[31] G. Sala, F. Haberl, J. José, A. Parikh, R. Longland, L. C. Pardo, and M. Andersen. Constraints on the mass and radius of the accreting neutron star in the rapid burster. *The Astrophysical Journal*, 752(2): 158, 2012. 28

[32] D. S. Sivia. *Data analysis: a Bayesian tutorial*. Oxford university press, 1996. 29

[33] A. K. Soper. Partial structure factors from disordered materials diffraction data: An approach using empirical potential structure refinementempirical potential monte carlo simulation of fluid structure. *Phys Rev. B.*, 72(10):104204, September 2005. 27

[34] A. Y. Toukmaji and J. A. Board Jr. Ewald summation techniques in perspective: a survey. *Comput. Phys. Commun.*, 95:73–92, 1996. 19

[35] J. D. van der Waals. Thermodynamische theorie der capillariteit onder de hypothese van een continue verandering van dichtheid. *Verhandelingen der Koninklijke Nederlandsche Akademie van Wetenschappen te Amsterdam*, 1:1–56, 1893. 13

[36] L. Verlet. Computer 'experiments' on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, 159(98):98–103, 1967. 11, 14

[37] A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol*, 103:227–249, 1976. 11

Chapter 3

Information theory applied to the study of the ordering in liquids

The determination of special molecular arrangements in disordered phases such as liquids (and plastic crystals) is inherently difficult due to its lack of periodicity as in the case of crystalline solids, section 2.2 presented a general method to study molecular liquids capable to unveil the details of the molecular ordering from small molecules to systems as big as a protein. This chapter shows a method to extract some general features of a liquid phase (or a plastic crystal) without going into such details. The method analyzes the probability distributions describing position and orientational molecular ordering within the framework of information theory.

The chapter is organized as follows: The first section introduces the main concepts of Information theory based on the pioneering article of Claude E. Shannon(13) of 1948 in a general framework. Those concepts are used to analyze the structure of liquids in the second section. This framework is applied to carbon tetrachloride (CCl_4), which is known for its tetrahedral symmetry that has been well studied (1, 3, 4, 8, 9, 11, 12). The results shown in this chapter have been published in the paper: **Characterizing ordering in liquids: An information theoretic approach**, Journal of

Non-Crystalline Solids, 407 (2015) (10).

3.1 Information theory

Information theory is considered to be born with the seminal paper of Claude E. Shannon in the early 50's (13). This theory was aimed to quantify the amount of information that a message carries and how can it be transmitted through a noisy channel without a significant loss of quality. However, this theory quickly crossed its borders and it is now applied to a vast number of fields such as message encryption (14), analysis of seismic data to search oil (2), and to decide which is the best strategy to follow in gambling games (5). In the context of this thesis, information theory will be used to quantify both the amount of information carried by a n-dimensional probability distribution and the correlation between two or more variables that characterize the short range order of a liquid. The basic notions of information theory are presented following the excellent paper of Matsuda et al. (7).

3.1.1 Shannon entropy

Given an n-dimensional probability distribution function (PDF) described by variables A_i, $i = 1, \dots n$ and that each variable A_i can take the discrete values a_i, being the number of these values for each variable A_i equal to $\tilde{N}_i$. The Shannon entropy associated to the normalized discrete probability distribution $p(a_1, \dots a_n)$ will be:

$$H(A_i) = -\sum_{\{a_i\}} p(a_i) \ln p(a_i). \tag{3.1}$$

The Shannon entropy so defined has an upper and a lower bound depending on the way the PDF has been generated. To study those boundings, the number of total bins of the n-dimensional distribution are assumed to be N (so that $N = \sum_{i=1,n} \tilde{N}_i$). On the one hand, if the probability distribution is flat, i.e. if all the values of probability are $p(a_1, \dots, a_n) = 1/N$, the PDF contains no information and in this case the associated Shannon entropy will be maximal and equal to ln N:

$$H(A_1 \dots A_n) = -\sum_{k=1}^{N} \frac{1}{N} \ln \frac{1}{N} = \ln N. \tag{3.2}$$

On the other hand, if the probability distribution has only one bin: i.e. one bin with probability one and all the rest with null probability, the Shannon entropy associated to this PDF will be zero. Therefore, the Shannon entropy can also be understood as a measurement of how "peaky" the landscape of a PDF is: smooth probability distributions will have high entropies. This fact is used, for example, in maximum entropy methods to find the smoothest probability distribution able to describe a data set.
An analogous definition can be used to calculate the Shannon entropy of any subset of variables $A_{1,\dots,m}$ where $m < n$, just marginalizing the probability distribution by the sum of the variables $A_{m+1,\dots,n}$ and defining thus the Shannon entropy of this reduced ensemble as:

$$p_{a_1,\dots,m} = \sum_{\{a_{m+1},\dots,n\}} p\left(a_{m+1}, \dots, n\right). \tag{3.3}$$

In this case the limits for the Shannon entropy hold, being in this case N the number of bins of the m-dimensional probability distribution.

3.1.2 Mutual information

Mutual information quantifies the correlation between two or more variables whose dependence is encoded in a certain PDF. For the sake of clarity, the two-dimensional case is described first and the case of higher dimensions is studied later.

3.1.2.1 The two-dimensional case

In a 2D probability distribution the mutual information is defined as:

$$\begin{aligned} I_2(A_1, A_2) &= \sum_{a_1,a_2} p(a_1, a_2)\, ln \frac{p(a_1, a_2)}{p(a_1)p(a_2)} \\ &= H(A_1) + H(A_2) - H(A_1 A_2). \end{aligned} \tag{3.4}$$

If variables A_1 and A_2 are independent, their joint probability can be rewritten as $p(a_1, a_2) = p(a_1) \cdot p(a_2)$, in this case the calculated mutual information is zero. On the other hand, the maximum value that I_2 can hold when the variables are fully correlated is the smallest value among the entropies of variables A_1 and A_2:

$$0 \leq I_2(A_1, A_2) \leq min\{H(A_1), H(A_2)\} \tag{3.5}$$

Figure 3.1 shows the cartoons of two-dimensional probability distributions ordered by increasing values of the Shannon entropy $H(A_1A_2)$ and zero mutual information $I(A_1, A_2)$. While figure 3.2 shows the case of an increasing Shannon entropy $H(A_1A_2)$ and increasing mutual information $I(A_1, A_2)$. It must be noted that the values of Shannon entropy and mutual information do not determine univocally the shape of a PDF. Thus, the figures are only intended to show how the shape of certain simple PDFs affects both their associated Shannon entropy and mutual information. To keep things as simple as possible, another restriction is added: the probability distributions will be generated by allowing their pixels to be only switched on or off, in other words pixels in "grey" are not allowed. That means that the calculations of the Shannon entropies are straightforward for both one and two-dimensional PDF's: $H(A_1) = H(A_2) = H(A_1A_2) = N_{on}$, where N_{on} is the number of pixels switched on.

For a 2-dimensional PDF with the same number of pixels per side N, the maximum number of switched on pixels for $p(a_1, a_2)$ will be N^2 and thus $H_{max}(A_1A_2) = 2\ln N$ and $H_{max}(A_1) = H_{max}(A_2) = \ln N$. Therefore, the maximum mutual information possible for such a PDF is simply $I_2^{max}(A_1, A_2) = \ln N$.

The simplest case is presented first: the variables of the probability distribution are not correlated, i.e., $I(A_1, A_2) = 0$ (see figure 3.1). Panel a) shows the only PDF with $H(A_1A_2) = I(A_1, A_2) = 0$: a probability distribution with a unique pixel switched on. For this case all the Shannon entropies are zero, and so also the mutual information. One way to increase the Shannon entropy keeping the mutual information equal to zero is sim-

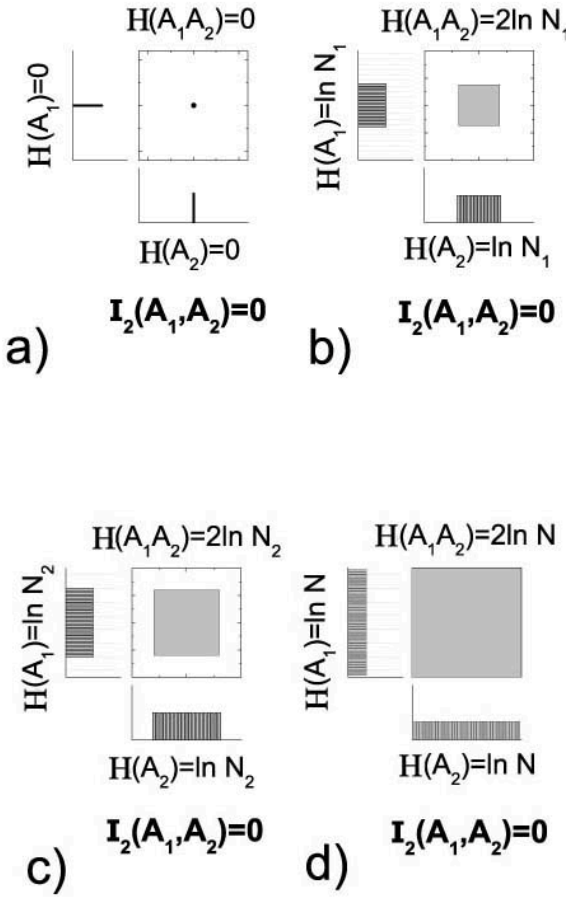

Figure 3.1: Two variable mutual information and Shannon entropy of some selected two-dimensional probability distributions. Distribution were chosen to keep zero the mutual information and an increasing value of the Shannon entropy. This figure includes the calculations of the Shannon entropies for the 2D probability distributions together with the calculations for the 1D projections and the mutual information calculations for all the cases following equation 3.4.

ply to switch on pixels keeping the squared shape of the PDF (and, in fact, keeping any other symmetric shape) until all pixels are on, being thus the Shannon entropy maximal.

Figure 3.2 shows another way to increase the Shannon entropy keeping

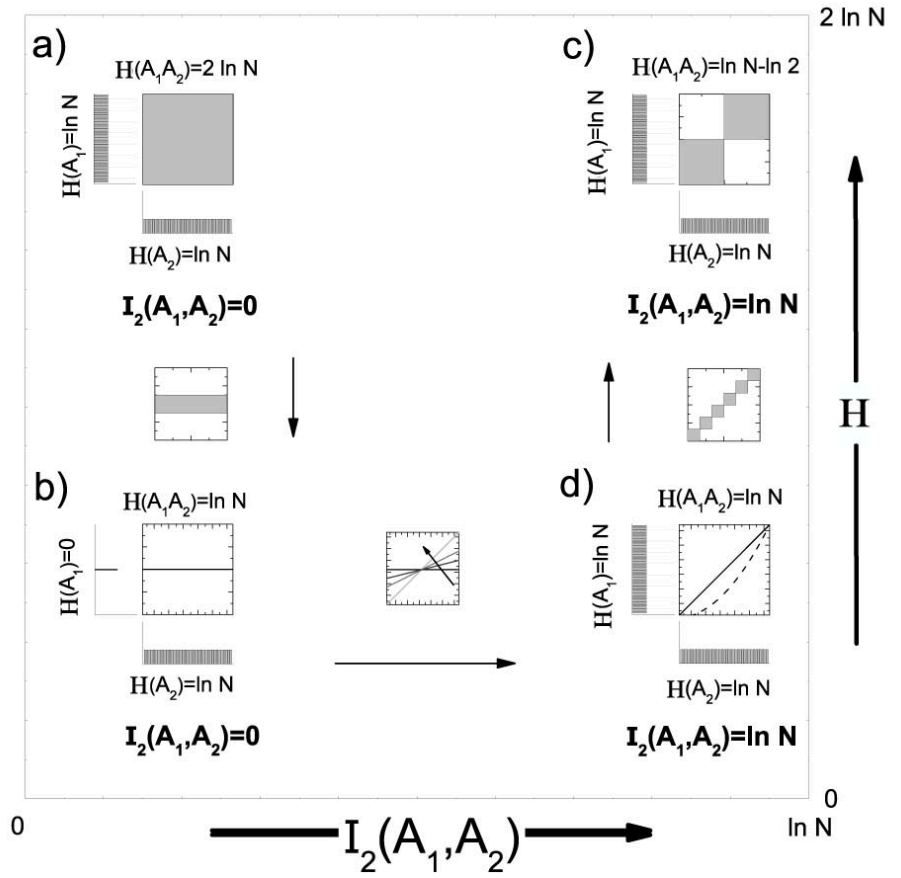

Figure 3.2: Examples of two-dimensional probability maps ordered by increasing values of Shannon entropy (ordinate) and mutual information (abscissa). The figure includes both 2D probability distributions together with their 1D projections and the calculations of their entropies. The inset between panels b) and c) indicates a 2D series of 2D distributions with increasing mutual information in the sense of the arrow. Panel c) shows two 2D distributions with a maximal mutual information: one with a linear correlation between variables (full line) and one with a non-linear correlation (dashed line).

the mutual information equal to zero. This is to switch on pixels forming a line perpendicular to any of the two variable A_1 or A_2. In this case, zero mutual information reflects the fact that the knowledge of the value of one variable does not help at all to determine the value of the other one. On the one hand, in order to increase the Shannon entropy keeping the mutual information equal to zero such a line can be made "thicker" by switching

on adjacent pixels.

On the other hand, to increase the mutual information by keeping the Shannon entropy constant, the slope of the horizontal line of figure 3.2 can be simply increased: the number of pixels switched on does not change and therefore $H(A_1A_2)$ remains constant and equal to $\ln N$. However the number of bins for the projected PDF's A_1 and A_2 start to grow and so it does their associated Shannon entropies. Maximum mutual information is reached when the line has a slope equal to one: in that case knowing A_1 completely determines the knowledge of A_2. It should be pointed out that the definition of mutual information is also capable to handle with the case when two variables are not linearly correlated. In this case (see the dashed line of figure 3.2), mutual information keeps being maximum. Therefore, in order to discover correlations between variables it is better to use the mutual information than the correlation function $\sigma_{A_1A_2} = \sum_{a_1,a_2} (a_1 - \bar{a_1}) \cdot (a_2 - \bar{a_2})$, this last is only able to describe linear correlation. In order to increase the Shannon entropy, keeping the mutual information maximal, the line describing the PDF has to be made thicker.Panel c) shows a PDF that has the aforementioned properties (both maximum Shannon entropy and mutual information).

3.1.2.2 The three-dimensional case

The mutual information for a probability distribution in three dimensions can be calculated having into account its relationships with the Shannon entropies of the three-dimensional PDF, and its projections in two and one dimensions (7):

$$\begin{aligned} I_3(A_1, A_2, A_3) = H(A_1) + H(A_2) + H(A_3) \\ -H(A_1A_2) - H(A_1A_3) - H(A_2A_3) \\ +H\left(A_1A_2A_3\right). \end{aligned} \tag{3.6}$$

Again, the mutual information associated to a three-dimensional probability distribution measures how well correlated are the three variables.

However, in this case the limits for the mutual information are very different than the ones of the two-dimensional case since the lower bound can be negative (7):

$$\begin{aligned} -\, min\left\{H(A_1), H(A_2), H(A_3)\right\} \\ \leq\ I_3(A_1, A_2, A_3) \leq \\ min\left\{H(A_1), H(A_2), H(A_3)\right\}. \end{aligned} \tag{3.7}$$

In order to clarify why the mutual information for a three-dimensional PDF can be negative, the example proposed by Matsuda et al. is used. It consists of a three spin system that can be up or down having an interaction among them ruled by a Heisenberg Hamiltonian: $\hat{H} = -J(x_1x_2 + x_2x_3 + x_1x_3)$ (do not confuse with the Shannon entropy H), where $J = \pm 1$. The case where variables x_i can take only the values ± 1 is studied. For this Hamiltonian the shape of the three-dimensional PDF will be very simple since it is formed by eight 3D pixels, called voxels from now on, whose state can be easily calculated using:

$$p\left(x_1, x_2, x_3\right) = \frac{e^{-\beta \hat{H}}}{Z} \tag{3.8}$$

where $\beta = 1/K_BT$ is related to the inverse of the temperature, and Z is the partition function that, for this case, is $Z = 2e^{3\beta J} + 6e^{-\beta J}$. The PDF for two extreme cases is calculated: when the temperature is infinity ($\beta = 0$) and when the temperature is zero ($\beta = \infty$). For the first case ($\beta = 0$), calculation in both cases $J = \pm 1$ is very simple since all voxels are switched on and therefore the Shannon entropy of the system is maximal and the mutual information is zero.

The interesting case is when the temperature is zero and therefore $\beta = \infty$. In this case very different PDF shapes for the cases $J = \pm 1$ appear. Panel a) of figure 3.3 shows the PDF for the case $J = 1$, parallel spins are giving the minimum energy of the system and therefore there is a well defined zero Kelvin state with all spins either up or down. In this case the three body mutual information of the system is positive $I_3(X_1, X_2, X_3) = ln2$. The three-dimensional PDF together with its projections in two and one

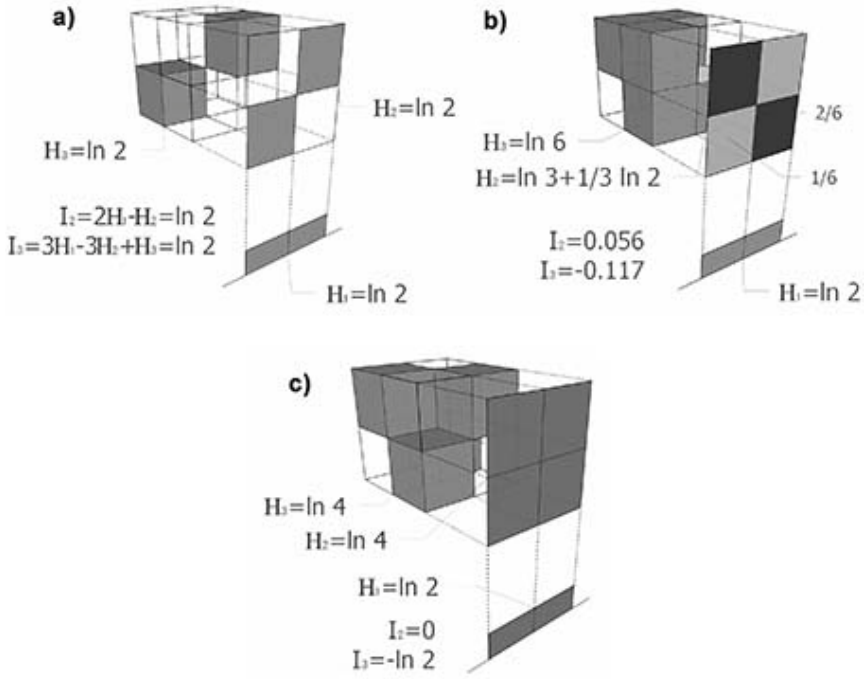

Figure 3.3: Symmetric three-dimensional probability distributions together with their two and one-dimensional projections, ordered by decreasing values of mutual information. Figures a and b are coming from a three spin system at zero temperature in the non-frustrated and frustrated case respectively. Figure c is one example of the maximum frustration that can be achieved from a three-dimensional probability distribution.

dimensions needed to calculate the mutual information are also shown in panel a) of figure 3.3. As it can readily be seen the 3D PDF is relating the variable one by one as in the 2D case, so that the dependency between variables is maximum. Moreover, the projections in two dimensions also allow to see that there is a relationship between variables, and in fact it fully determines their dependence. Roughly speaking, the 3D PDF is well behaved and its 2D projections are giving the information concerning the dependence between variables.

The case where $J = -1$ that favors anti-parallel spin interaction is now studied: this is a well known case of a geometric frustrated system. Panel b) of figure 3.3 shows the three-dimensional PDF for this system when

$\beta = \infty$. For this case the mutual information is negative and equal to $I_3(X_1, X_2, X_3) = -0.117$. As it can readily be seen the three-dimensional PDF is not as well behaved as in the last case: if a spin is up, i.e. if a cut is performed on the 3D PDF and only the upper voxels are taken, there are a lot of possibilities of arranging the other two spins with equal probability. Moreover, the projections in two dimensions do not make possible to deduce the shape of the three-dimensional PDF. Because of this, the Shannon entropies of the two-dimensional projections of the original 3D PDF have a high value and, since they are subtracting terms in the definition of the three variable mutual information 3.6, the total value is negative.

The main point that should be stressed out with this very simple example is that a negative mutual information implies a correlation between three variables that can not be seen as the sum of a pairwise dependence of two variables. In other words, a careful look at the 3D probability distribution is necessary to get any information about how variables are correlated, since its two-dimensional projections have all the information messed up resulting in a high entropic PDF. Following this idea, panel c) of figure 3.3 shows a PDF with the maximum negative mutual information: its two-dimensional projections have a maximum Shannon entropy and their mutual information is zero, i.e. little about the 3D PDF can be inferred from the 2D projections. However the PDF has a well defined shape in three dimensions, and thus the mutual information is negative and maximal ($I_3(X_1, X_2, X_3) = -\ln 2$).

Mutual information also allows the Shannon entropy of an n-dimensional PDF to be written as the sum of the Shannon entropies of the one-dimensional projection of the PDF plus mutual information terms:

$$H_{tot} = \sum_{i=1}^{n} H_i(A_i) - \sum_{i<j} I_2(A_i, A_j) + \sum_{i<j<k} I_3(A_i, A_j, A_k) + \cdots \tag{3.9}$$

This relationship makes possible to calculate the Shannon entropy of

a high-dimensional PDF using lower dimensional PDFs (being sometimes *high-dimensional* as low as $n \geq 4$). This is interesting because the number of n-dimensional voxels increases as N^n which leads to two problems. First, since the memory of the computer and its speed is finite it will not be easy to handle n-dimensional matrices for a reasonable number of voxels N. Second, calculations of PDFs are usually done by counting events and then normalizing them by their total amount. The number of events that must be counted to have a reasonable PDF increases with its dimensionality (in fact it increases as N^n). If the number of events is not high enough the PDF will consist in series of sparse switched on voxels making any calculation of the Shannon entropy or mutual information meaningless.

3.2 Information theory applied to the study of the liquid structure

This section shows the development of a framework to study molecular liquids without entering in the detailed analysis required by the Euler angles, so it can capture the general features. This can be done using the information theory approach to study the six-dimensional probability distribution function (PDF) (see section 2.2) describing the relative position and orientation of a molecule at a distance r from the central one $g\left(r, \Omega_{pos}, \Omega_{or}\right)$, where Ω_{pos} are the angles θ_{pos} and ϕ_{pos} describing the position of the center of mass of a neighbor molecule with respect to the central one at a distance r and Ω_{or} are the three Euler angles θ_{or}, ϕ_{or} and ψ_{or} describing the relative molecular orientation.

The example of carbon tetrachloride is used to perform the analysis on the 6D-PDF $g\left(r, \Omega_{pos}, \Omega_{or}\right)$ obtained in subsection 2.2.2, the details of the MD simulation and the Euler angles analyses can be found there.

3.2.1 Liquid ordering and mutual information

If only the first three terms of equation 3.9 are used, the positional contribution to the total entropy at a given distance can be defined as follows:

$$S_{pos}(r) = S(\cos(\theta_{pos})|r) + S(\phi_{pos}|r) - I\left(\cos(\theta_{pos}), \phi_{pos}|r\right) \tag{3.10}$$

It must be pointed out that the calculated entropy from $g\left(r, \Omega_{pos}, \Omega_{or}\right)$ is related to the thermodynamic entropy by an approximation using two-body molecular interactions (6). This is why the notation is changed from H to S.

The entropy scaled to its asymptotic value for long distances (S_{pos}/S^{∞}_{pos}) and the mutual information of the PDF related to the positional variables $\cos(\theta_{pos})$ and ϕ_{pos} are shown in the panel a) of figure 3.4. As expected, for long distances mutual information tends to zero and entropy to a fixed and maximal value. Figure 3.4 also shows that both mutual information and entropy are correlated, so that when one increases the other decreases. In order to investigate this fact the inset of the figure shows a different rescaling for both magnitudes so that they have a limited range variation from zero to one ($\xi^{resc} = \xi - \xi_{min}/\xi_{max} - \xi_{min}$, being ξ either the entropy or the mutual information), it shows that they are completely correlated so that $S^{resc} = -I^{resc}$. This is due to the small differences between the values of entropies for short and long distances coming from the 1D projections of $g(\cos(\theta_{pos}), \phi_{pos})$, i.e. $S(\cos(\theta_{pos}))$ and $S(\phi_{pos})$.

Panel b) of figure 3.4 shows the partial radial distribution of the carbon atoms of CCl_4 $g_{CC}(MCN)$, i.e. the center of mass of the molecule. The mean coordination number (MCN) is used instead of the distance. Comparing both panels a) and b) it can be seen that for long distances minima and maxima of $g_{CC}(MCN)$ are correlated with maxima (minima) of the mutual information (entropy) from the probability distribution describing the position of the molecular centers of mass. In order to help the comparison, the figure also shows $\|g_{CC}(MCN) - 1\| + 1$ to flip up the minima of $g_{CC}(MCN)$. This figures shows that the liquid does not get monotonically disordered with the distance.

In fact, there are some distance intervals where the liquid is more ordered. Figure 3.5 shows the 2D-PDF describing the position of the centers

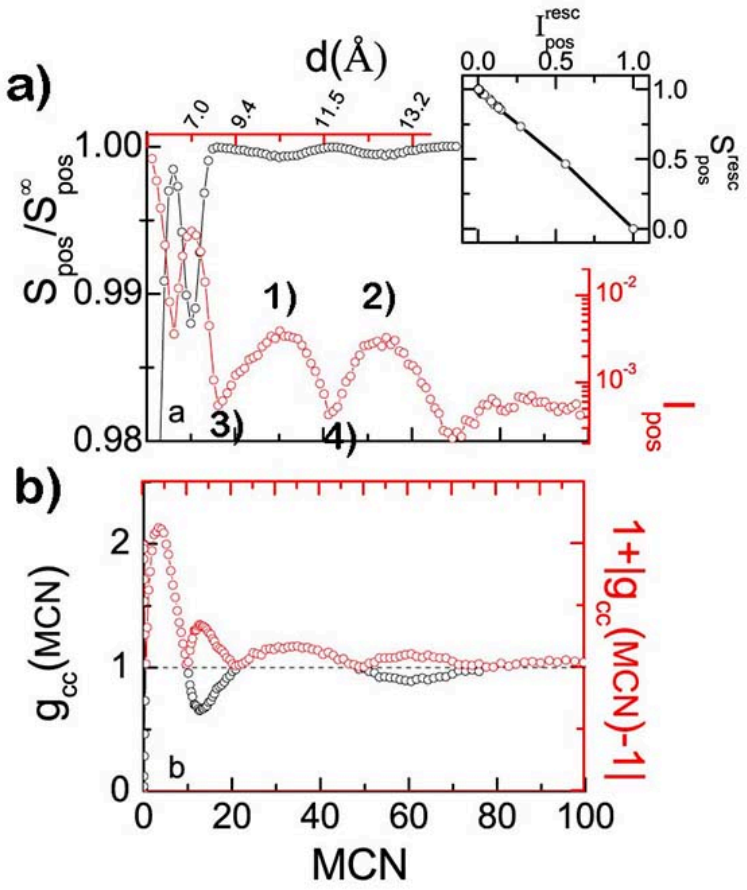

Figure 3.4: Panel a) shows both the rescaled entropy and the mutual information for the 2D probability distributions describing the molecular position such as that of panel (a1) of figure 3.5 as a function of the molecular coordination number (MCN). Panel b) shows the radial distribution function $g_{CC}(MCN)$ together with its flipped version $||g_{CC}(MCN) - 1|| + 1$

of mass $g(cos(\theta_{pos}), \phi_{pos})$ for maxima and minima of the mutual information selected in 3.4. From these figures it is clear that minima of mutual information (maxima of entropy) is related to regions where there is not a particular ordering of the centers of mass of neighboring molecules (regions with flat distributions). On the contrary around the maxima of mutual information (minima of entropy) it is clear that the liquid is more ordered.

The positional contribution to the total entropy has been analyzed to

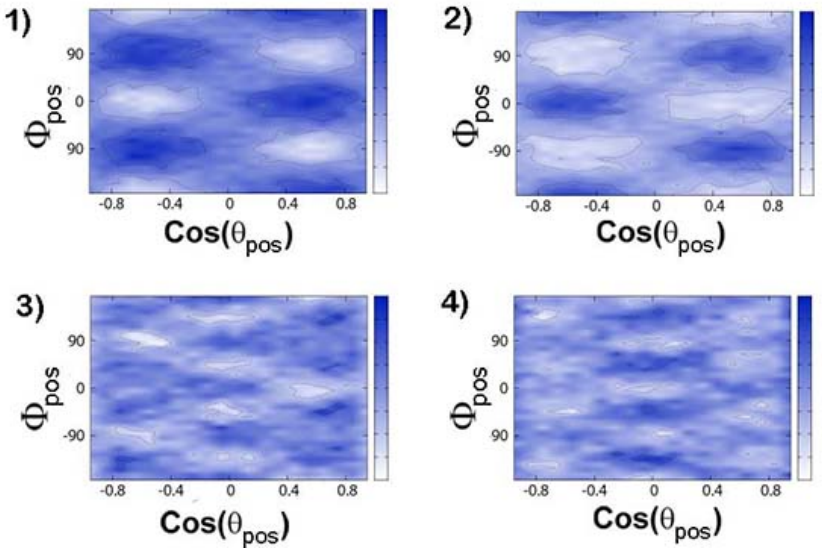

Figure 3.5: The 2D-PDF describing the position of the centers of mass $g(cos(\theta_{pos}), \phi_{pos})$ for maxima and minima of the mutual information marked in Figure 3.4

introduce the information theory in the context of this thesis, namely the analysis of the molecular ordering in liquids. The same procedure applied in this section can be used to study the orientational contributions by analyzing the Euler angles $(\theta_{ori}, \phi_{ori})$ and ψ_{ori}.

3.3 Bibliography

[1] P. Egelstaff, D. Page, and J. Powles. Orientational correlations in molecular liquids by neutron scattering carbon tetrachloride and germanium tetrabromide. *Molecular Physics*, 20(5):881–894, 1971. 37

[2] P. E. Haggerty. The corporation and innovation. *Strategic Management Journal*, 2(2):97–118, 1981. 38

[3] P. Jedlovszky. Structural study of liquid methylene chloride with reverse monte carlo simulation. *The Journal of chemical physics*, 107(2): 562–570, 1997. 37

[4] P. Jóvári, G. Mészáros, L. Pusztai, and E. Sváb. The structure of liquid

tetrachlorides ccl4, sicl4, gecl4, ticl4, vcl4, and sncl4. *The Journal of Chemical Physics*, 114(18):8082–8090, 2001. 37

[5] J. L. Kelly Jr. A new interpretation of information rate. *Information Theory, IRE Transactions on*, 2(3):185–189, 1956. 38

[6] T. Lazaridis and M. Karplus. *J. Chem. Phys.*, 105:4294, 1996. 48

[7] H. Matsuda. *Phys. Rev. E.*, 62:3096–3102, 2000. 38, 43, 44

[8] A. Narten, M. Danford, and H. Levy. Structure and intermolecular potential of liquid carbon tetrachloride derived from x-ray diffraction data. *The Journal of Chemical Physics*, 46(12):4875–4880, 1967. 37

[9] L. Pardo, N. Veglio, F. Bermejo, J. L. Tamarit, and G. Cuello. Experimental assessment of the extent of orientational short-range order in liquids. *Physical Review B*, 72(1):014206, 2005. 37

[10] L. C. Pardo, A. Henao, and A. Vispa. Characterizing ordering in liquids: An information theoretic approach. *Journal of Non-Crystalline Solids*, 407:220–227, 2015. 38

[11] S. Pothoczki, L. Temleitner, P. Jóvári, S. Kohara, and L. Pusztai. Nanometer range correlations between molecular orientations in liquids of molecules with perfect tetrahedral shape: Ccl4, sicl4, gecl4, and sncl4. *The Journal of chemical physics*, 130(6):064503, 2009. 37

[12] R. Rey. Quantitative characterization of orientational order in liquid carbon tetrachloride. *The Journal of chemical physics*, 126(16):164506, 2007. 37

[13] C. E. Shannon. A mathematical theory of communication. *Bell system technical journal*, 27:379–243, 1948. 37, 38

[14] C. E. Shannon. Communication theory of secrecy systems*. *Bell system technical journal*, 28(4):656–715, 1949. 38

Chapter 4

A first study: liquid and plastic hexachloroethane

Previous chapters introduced the methods and theory frameworks in order to investigate the structure of disordered phases. This chapter shows the results for a system with a plastic phase and the comparison to its liquid phase using ANGULA and Information theory. The system chosen was the hexachloroethane molecule (C_2Cl_6). Some of the results shown in this chapter, the results on the molecular structure C_2Cl_6 and the results on the liquid structure of C_2Cl_6 have been published in the papers: **Insights into the determination of molecular structure from diffraction data using a Bayesian algorithm** (10). and **Competing structures within the first shell of liquid C_2Cl_6: A molecular dynamics study** (11).

This chapter is organized as follows: an introduction about the different phases of hexachloroethane and the work done in the plastic phase is presented. The computational details are given and next the obtention of the intra molecular parameters of the molecule from neutron diffraction is shown. The plastic phase simulation is compared to experiments and previous works. Finally, the short range and long range ordering of the liquid and plastic crystal phases is compared.

4.1 Motivation

Hexachloroethane (C_2Cl_6) is an interesting material because of its molecular symmetry, which allows it to have a plastic crystal phase. Figure 4.1 shows a cartoon of the molecule, it is an ethane molecule. C_2Cl_6 can be considered as a nonassociated liquid (it is well described by the van der Waals picture). This means that the steric effects are the most important to understand its structure. C_2Cl_6 has three different solid phases (1, 15) (orthorombic between 4 and 318K, monoclinic between 318 and 344K and a plastic phase stable from 344K up to the melting temperature $T_m = 458K$ (23)). This plastic phase has a body centered cubic unit (bcc). The positional disorder in the plastic phase come from the fact that the space group is Im3m and a corresponding site group m3m. But the molecule has a lower symmetry $\overline{3}m$. Therefore, the orientational disorder appears for the molecule to fulfill with the site group symmetry (6).

The intramolecular geometry of C_2Cl_6 has been investigated before by means of electron diffraction in the gaseous and solid states (14, 25). In those works, the authors obtained the bond and angles parameters that define the geometry of the C_2Cl_6, but those parameters were not able to reproduce the structure of the molecule obtained from neutron diffraction experiments in the plastic phase. Therefore, the intramolecular geometry of the molecule was studied using a Bayesian fit to the experiments. This will produce a better set of molecular parameters to use as an input in MD, with a better agreement with the plastic phase.

Regarding the structure of the plastic phase, in 1981 Gerlach et.al (6) conducted a neutron powder and Single-Crystal investigation and later Gerlach et.al (7) compared the previous results to a model structure of the plastic phase from a Monte Carlo (MC) method. In 1988 Gerlach and Prandl (5) studied the orientational ordering in the plastic phase using diffuse X-ray scattering and later (Gerlach et al. (8)) with elastic and quasielastic neutron scattering. Criado and Muñoz (1) used MD to simulate the plastic phase of C_2Cl_6 and they used the results to interpret an inelastic neutron scattering experiment (15). Negrier et al. also studied the symmetry of the plastic phase have (16). It will be shown a comparison of the MD results with MD data from Criado and Muñoz and also with the experiments on the

plastic phase from Negrier et al. and those from Gerlach et.al. No experimental studies on the structure of liquid C_2Cl_6 have been performed (to our knowledge). Given its van der Waals symmetry (as explored by Slovokhotov et al. (24)), the boiling temperature is 1K above the melting point under ambient pressure making C_2Cl_6 to sublime. Therefore a simulation study is ideal to gain knowledge on this liquid to obtain further comparisons to the plastic phase.

4.2 Computational details

4.2.1 Simulation details

MD simulations were performed for the plastic and liquid phases of C_2Cl_6 on a 2000 molecules system. In the case of the bcc plastic phase, a supercell of 10x10x10 unit cells was used with the positions of the molecules corresponding to the bcc crystal and a random orientation for all the molecules and the simulation was performed at $T = 400K$ and $P = 1atm$. In order to obtain the liquid phase, the equilibrated final configuration from the plastic crystal was used for a simulation at the thermodynamics conditions of the liquid (22), namely T = 458K and P = 1atm. Simulations were also performed on a small system of 504 molecules using the bcc lattice (6x6x7 unit cells) at 6 different temperatures (within the range of the bcc plastic phase) to have results on the lattice parameter expansion, T = 344, 360, 380, 400, 413, 423K. For all the simulations, 1000 configurations were collected after equilibration and saved every 1ps.

All the MD simulations were performed using the Gromacs 4.5 (12) package. The potential parameters were chosen from the Gromos53a6 (17) force field (6-12 Lennard Jones). Other general conditions were: Δt = 2fs (during 500ps), shifted cut-off from 16 to 17Å for Lennard-Jones interactions and 20Å for coulomb pairs. The Particle Mesh Ewald (PME) was used method beyond the electrostatic cut-off for the reciprocal space sum. Finally the analysis was made over the last uncorrelated 300ps.

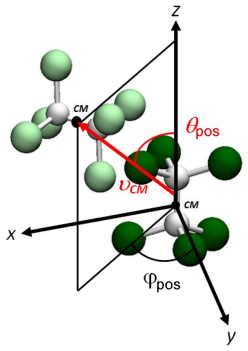

Figure 4.1: Reference axis definition for the Euler angles study on the molecule C_2Cl_6. Angles θ_{pos} and ϕ_{pos} describe the position of any molecule relative to a reference molecule (with darker chlorides in the figure) using the vector v_{CM} that joins their centers of masses.

4.2.2 Axes definition

A rigid molecule model was used and the internal molecular motions were neglected as this molecule does not have conformers. This means that all the degrees of freedom of the molecule can be known by studying its position and orientation. To do that, the proper Euler angles convention (9) introduced in section 2.2 is used. The C_2Cl_6 molecule defines a coordinate system as shown in Figure 4.1). One remark from the angle definitions is that θ_{ori} contain information about the relative orientation of the C-C bonds, while ϕ_{ori} and Ψ_{ori} describe the orientation of the chloride shells.

4.3 The molecular structure of hexachloroethane

The knowledge of the molecular geometry, angles and bond length are a necessary input for the MD simulation of the rigid molecule. This section shows the obtention of these parameters by using a bayesian fit to neutron diffraction experiments (see section 2.3). This has been done to improve the agreement with the experimental data compared to the intramolecular geometry of C_2Cl_6 given by other authors (14, 25). Their parameters are based on both the gaseous and solid states. The fitting presented here corresponds with the plastic phase that is of our interest and therefore a

better description of the‘plastic phase is expected.

4.3.1 Experimental

A sample of C_2Cl_6 with a 99% purity was purchased from Sigma-Aldrich and used to perform the diffraction experiments without any further purification. The experiment was performed in the plastic phase of C_2Cl_6 at a temperature of 400K . The diffractograms were obtained at the liquids and glasses neutron diffractometer D4c at the Institute Laue-Langevin (Grenoble, France) (3), using a wavelength of $\lambda = 0.5$ Å and an angular range for the detectors that yielded a scattering vector up to $q_{\max} \approx 23$ Å^{-1}. In order to correct and normalise the data, the empty cryostat, an empty sample holder, a boron powder sample and a vanadium rod were also measured, as in previous works (18). Absorption and multiple scattering corrections and normalization of data were performed using the program CORRECT (13). Additionally, inelastic corrections were also carried out by substracting a polynomial expansion in powers of q^2 (19).

4.3.2 Calculation of the intramolecular S(q) and RDF for C_2Cl_6

The expressions for the analytical intramolecular structure factor and radial distribution function (RDF) based on the knowledge of the geometry of the molecule is needed. In figure 4.2, a scheme of the molecular geometry illustrates the parameters that have been fitted to fully account for the molecular structure. Prior knowledge on the symmetry of the molecule allows to use the minimum number of parameters. In the case of ignoring completely the structure of the molecule, all the possible parameters should be fitted.

In order to use the minimum number of parameters and increase the robustness of the fit, the structure of the molecule is determined using only four independent parameters. As it can be seen in figure 4.2, there are five distinct distances, with two of them corresponding to C-C and C-Cl bond lengths (d_{CC} and d_{CCl}, respectively), and two different angles, α ($\angle$C-C-Cl) and the dihedral angle γ, which is responsible for the torsion within the

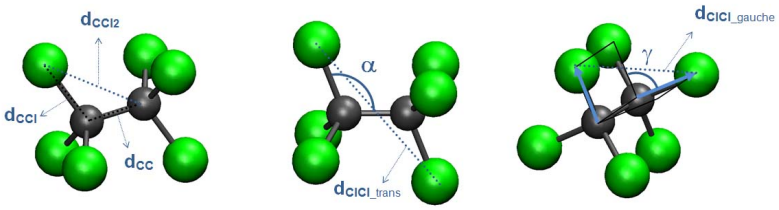

Figure 4.2: C_2Cl_6 molecule together with the parameters used to fit its structure factor.

molecule due to the rotation of the two chloride shells.

Taking the above two bond length distances and the angles α and γ as the independent parameters, and determine the rest of the distances:

$$d^2_{CCl2} = d^2_{CC} + d^2_{CCl1} - 2d_{CC}d_{CCl1}\cos(\alpha) \tag{4.1}$$

$$d^2_{Cl1Cl1} = 2(d_{CCl1}\sin(\alpha))^2(1-\cos(2\pi/3)) \tag{4.2}$$

The angle θ ($\angle$Cl-C-Cl) can now be calculated using the cosinus theorem:

$$\theta = \arccos\left(\frac{2d^2_{CCl1} - d^2_{Cl1Cl1}}{2d^2_{CCl1}}\right) \tag{4.3}$$

And to calculate the distances between chlorides bonded to different carbon atoms:

$$\begin{aligned} d^2_{ClCl_{trans}} = \quad & [-2d_{CCl1}\cos(\alpha) + d_{CC}]^2 \\ & +[d_{CCl1}\sin(\alpha) + (d_{CCl1}\sin(\alpha)\cos(\gamma-\pi))]^2 \\ & +[d_{CCl1}\sin(\alpha)\sin(\gamma-\pi)]^2 \end{aligned} \tag{4.4}$$

$$\begin{aligned} d^2_{ClCl_{gauche}} \quad & = [-2d_{CCl1}\cos(\alpha) + d_{CC}]^2 \\ & +[d_{CCl1}\sin(\alpha) + (d_{CCl1}\sin(\alpha)\cos(\gamma-\pi/3))]^2 \\ & +[d_{CCl1}\sin(\alpha)\sin(\gamma-\pi/3)]^2 \end{aligned} \tag{4.5}$$

The two parameter distances (d_{CC} and d_{CCl1}), together with the calcu-

lated ones (d_{CCl2}, d_{Cl1Cl1}, $d_{ClCl_{trans}}$, $d_{ClCl_{gauche}}$), will be needed to calculate the structure factor and the RDF. To estimate the contribution of each intramolecular distance to both functions, the following equations can be used (4, 20):

$$S_i(q) = \frac{\sin(qd_i)}{(qd_i)} \exp\left[-\frac{1}{2}(l_i q)^2\right] \tag{4.6}$$

$$\mathrm{RDF}_i(r) = \frac{1}{l_i} \exp\left[-\frac{1}{2}\frac{(r-d_i)^2}{l_i^2}\right] \tag{4.7}$$

where d_i is every particular distance i being considered, and l_i is the parameter in the Debye-Waller factor related to the distance i. The Gaussian approximation has been used in the expression of the RDF contributions (21).

Finally, each term has a different weight in the total structure factor given by:

$$\begin{aligned} S(q) \quad &= \{b_C b_C S_{CC}(q) + 6 b_C b_{Cl1} S_{CCl1}(q) \\ &+6 b_C b_{Cl2} S_{CCl2}(q) + b_{Cl1} b_{Cl1}[6 S_{Cl1Cl1}(q) \\ &+3 S_{ClCltrans}(q) + 6 S_{ClClgauche}(q)] \ / W_{intra}\} \end{aligned} \tag{4.8}$$

And for the radial distribution function:

$$\begin{aligned} \mathrm{RDF}(r) \quad &= \{b_C b_C \mathrm{RDF}_{CC}(r) + 6 b_C b_{Cl1} \mathrm{RDF}_{CCl1}(r) \\ &+6 b_C b_{Cl2} \mathrm{RDF}_{CCl2}(r) + b_{Cl1} b_{Cl1}[6 \mathrm{RDF}_{Cl1Cl1}(r) \\ &+3 \mathrm{RDF}_{ClCltrans}(r) + 6 \mathrm{RDF}_{ClClgauche}(r)]\}/W_{intra} \end{aligned} \tag{4.9}$$

where the numerical factors account for the number of possible pairs of two species, and $b_C = 0.646$ Å, $b_{Cl} = 0.9577$ Å are the coherent neutron scattering lengths of the two atom types of our molecule. The weights are used to normalize the distributions:

$$W_{intra}(r) = b_C b_C + 12 b_C b_{Cl} + 15 b_{Cl} b_{Cl} \tag{4.10}$$

4.3.3 Fitting procedure and results

The molecular structure of C_2Cl_6 has been fitted to diffraction measurements using the FABADA program. The RDF has been determined using the NS interval and once the fit has finished, a simulation annealing has been performed to ensure that the fit has arrived to the global minimum. The priors that have been used to perform the fit are Jeffrey's prior for all the length parameters (l and d) and a sin prior for angle γ since this is a polar angle. This priors were implemented tuning the jumps to be performed in a logarithmic scale for the lengths and in cosinus for the angle γ. Moreover, the results of a previous determination of the C_2Cl_6 structure (25) were included as priors.

Besides, the fits have been performed in both direct and reciprocal spaces, since in this way the determination of the molecular parameters is more robust (20).

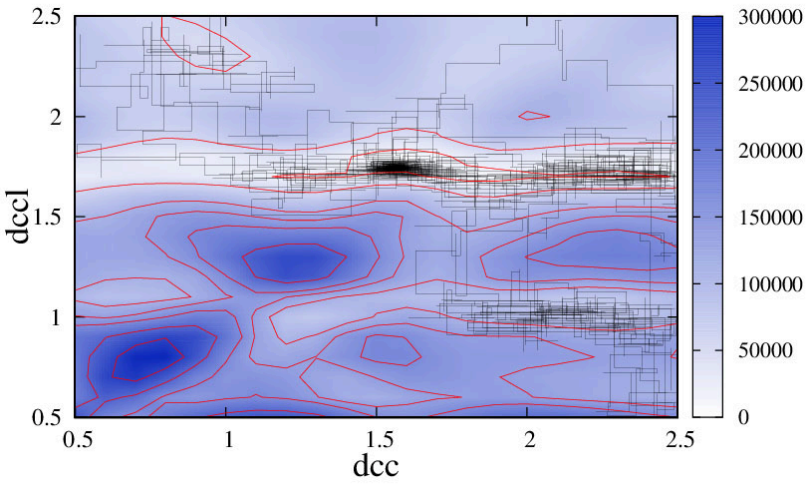

Figure 4.3: χ^2 of the function describing the molecular structure of C_2Cl_6 (Z axis) as a function of d_{CC} and d_{CCl} (see figure 4.2 for a scheme of the geometric parameters). A lighter shade means that a better fit is obtained for that parameter combination. The solid black line shows the path followed by the algorithm to jump from the local minimum at $(d_{CC}, d_{CCl}) = (2, 1)$ to the global minimum.

A complex bidimensional χ^2 landscape instance is shown in figure 4.3. The figure focuses on the region of the χ^2 landscape with parameter values between the ranges $(0.5 < d_{CC} < 1.5)$ and $(0.5 < d_{CCl} < 1.5)$. The solid black line depicts the path that the algorithm has followed during the process. Although plenty of the accepted parameter values belong to the global χ^2 minimum, the algorithm has been also able to move uphill and thus to travel across higher regions to reach other minima.

The experimental structure factor and RDF, together with their best fits, have been plotted in figure 4.4. The parameters related to the molecular structure are d_{CC}, d_{CCl}, α, and γ, the rest are related to intramolecular dynamics. A structure factor and RDF have also been generated using the

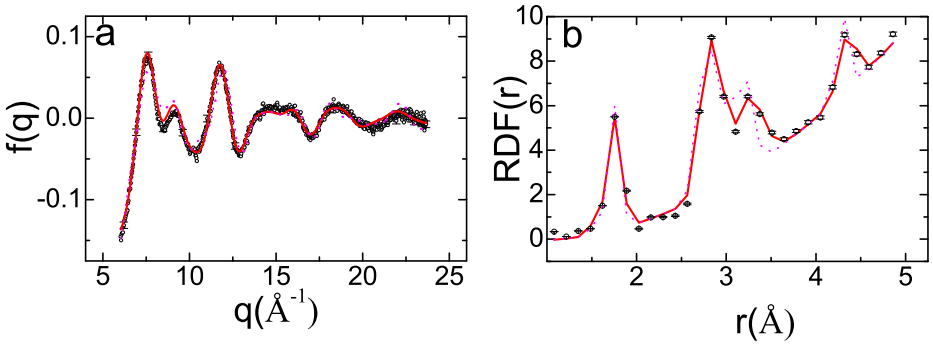

Figure 4.4: (a) Fit of the structure factor, and (b) the RDF using FABADA. Dots represent the experimental data (in reciprocal space (a) and their Fourier transform using NS sampling (b)), dashed lines represent the functions calculated with the parameters proposed in (25) and thicked lines show fits of the present work.

parameters taken from (25) (see figure 4.4). As it can be seen, those parameters don't describe as good as the generated ones the measured structure factor. After the concurrent fit in real and reciprocal spaces, a better description of the data was obtained, and the fits in both spaces are within data errors. A summary of the parameters obtained from the fit can be found in Table 4.1.

Since all parameter PDFs (except l_{CC}) could be reasonably well described by a Gaussian distribution, the error of the parameters is given as

the standard error σ of their normal distributions, and therefore containing a probability mass of a 67%. The γ value near 60° shows that the staggered conformation is the most stable.

Parameter	This work (Å)	Previous (25) (Å)
d_{CC}	1.590 ± 0.005	1.57 ± 0.06
d_{CCl1}	1.7670 ± 0.0009	1.74 ± 0.01
l_{CC}	0.137 ± 0.005	0.02 assumed
l_{CCl1}	0.0713 ± 0.0008	< 0.03
l_{CCl2}	0.088 ± 0.002	0.03 ± 0.02
l_{Cl1Cl1}	0.095 ± 0.002	-
$l_{ClCl_{trans}}$	0.152 ± 0.002	-
$l_{ClCl_{gauche}}$	0.098 ± 0.003	-
Angle	This work (°)	Previous (25) (°)
α	109.6 ± 1.2	109 ± 0.5
γ	60.2 ± 0.4	-

Table 4.1: Parameters obtained with FABADA compared to a previous work (25)

4.4 Neutron diffraction analysis of the plastic phase of hexachloroethane

In this section, attention is paid to the portion of the structure factor that contains the information on the intermolecular arrangements.

The lattice parameter expansion as a function of temperature was obtained (within the range of the plastic bcc phase of C_2Cl_6 $(344 - 457K)$) from the equilibrium configurations of the NPT ensemble and it was compared to previous experiments (16) and with other works: the MD simulation of Criado & Muñoz (1) and the powder neutron diffraction data of Gerlach et al. (6) in Figure 4.5. MD points compare well to the experimental ones and they fit better than the results of Criado & Muñoz.(1). This shows that the force field used in this work reproduces better the density and thermal expansion of the lattice than previous works. Not only different parameter were used, but also a different shape of potential: A

6-12 Lennard Jones was used instead of the Buckingham potential used by Criado & Muñoz.

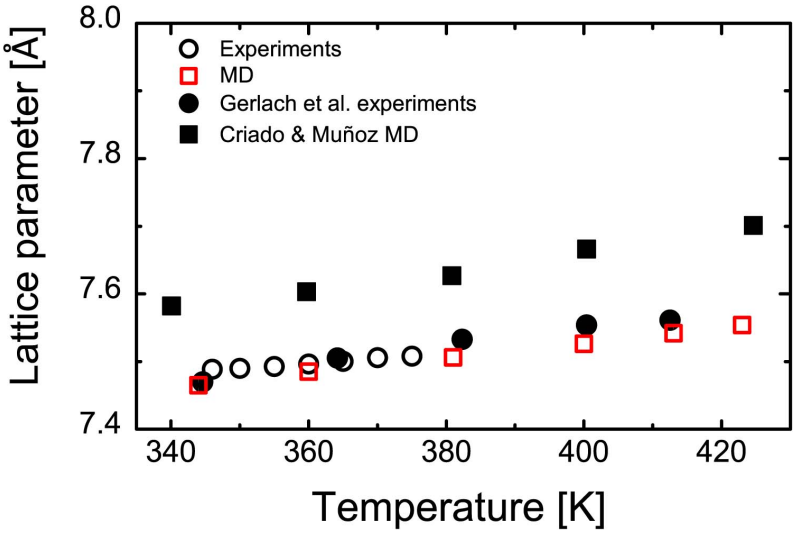

Figure 4.5: . The lattice parameter as a function of temperature within the plastic phase. Circles show experimental data, open symbols: our results and filled symbols: Gerlach et al. (6) Squares show MD data, again open symbols: our results and filled symbols: Criado & Muñoz (1)

The static structure factor S(q) was computed next as a fourier transform of the total RDF $g_{total}(r)$ weighted by the contribution of the atoms and their neutron scattering lengths ($b_c = 0.646$ and $b_c l = 0.9577$):

$$S(q) - 1 = 4\pi\rho \int_o^{r_c} (g_{total}(r) - 1) \frac{\sin qr}{qr} dr \tag{4.11}$$

it was chosen $\rho = 1$ as it is a scaling factor when compare with the experimental results and r_c is the cut off of the integration, given for a value with S(q) close to 0 and with small derivative for large r to avoid termination effects.

Figure 4.6 shows the comparison with the experimental S(q), it shows the q region up to $8 Å^{-1}$, beyond this point the contribution of the short

range ordering is small compared to the contributions from the internal geometry of the molecule.

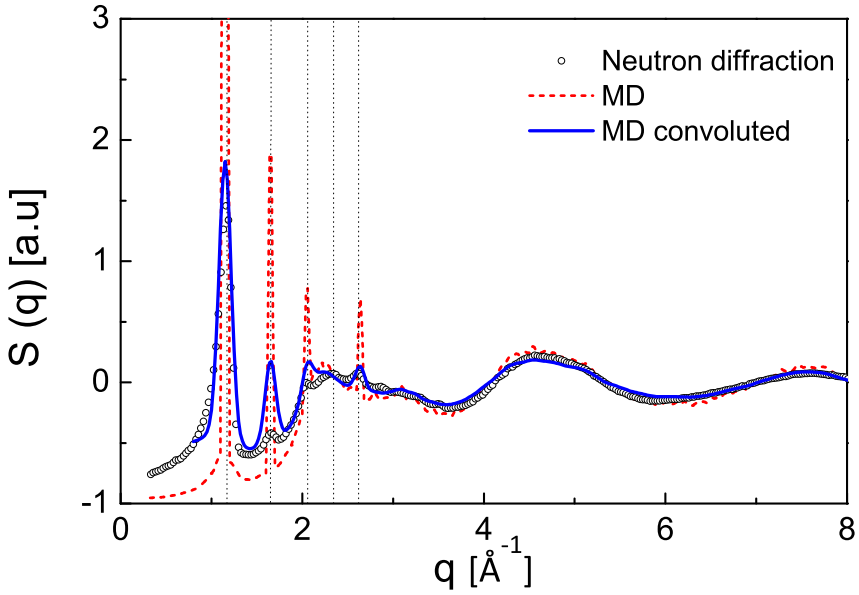

Figure 4.6: Static structure factor S(q) of C_2Cl_6 at $T = 400K$. Dots: experimental data. Dashed line: MD S(q) directly from trajectories. Thick line: MD S(q) convolved with instrumental resolution.

In Figure 4.6 the experimental $S_{exp}(q)$ is plotted with dots and the MD $S_{MD}(q)$ with a dashed line. Although there is not a way to completely separate the intramolecular contributions, the main contribution to the small q range is the inter molecular structure (20). This can be seen clearly in the $S_{MD}(q)$, where the defined peaks correspond to the positional ordering of the bcc phase. In order to support the last statement, the distances for the bcc crystal unit cell were computed using the equilibrium lattice parameter from MD: 7.54Å. Those distances are shown as dotted lines and fit with the MD peaks.

The function $S_{MD}(q)$ is more peaky than the experimental. In experiments, the broadening of the diffraction peaks come from two sources: the first is the one due to the disorder of the plastic phase, orientationally and

from thermal agitation. The second factor obeys to the limited instrumental resolution of the technique. The MD simulation intrinsically considers the first factor, but the second one is not present in simulations. In order to include this, it was performed a convolution of the $S_{MD}(q)$ with the instrumental resolution function. The details on this calculation were given in section 2.4. The result of that convolution corresponds to the thick line in Figure 4.6. Including the resolution function in the $S_{MD}(q)$ improves the agreement with the experiment, the intensity of the MD peaks decrease as a consequence of having into account the broadening of the peaks. The result shows that the instrumental resolution function should always be included in the $S_{MD}(q)$.

4.5 Short range order in the liquid and plastic phases

To the best of the knowledge of the author, there are no previous reports on the simulation of the liquid state of C_2Cl_6. Therefore, the results previously shown on the plastic phase present a validation for the force field to be used in the liquid phase. The final configuration of the bcc plastic phase was melted at the temperature of the liquid, T = 458K. The Mean-square displacement (MSD) is shown in Figure 4.7, a negligible positional motion (only thermal agitation) is seen for the plastic phase (thick line) as for the liquid phase (dotted line) the characteristic MSD slope shows the diffusion of the molecules, a fluid like property.

In order to characterize the dynamics of the liquid and plastic, the reorientational correlation function (RCF) was computed as $RCF_l = \langle P_l[\mathbf{U}_\alpha(0) \cdot \mathbf{U}_\alpha(t)]\rangle$. Where P_l is the first or second order Legendre polynomial, and the vectors $\mathbf{U}_\alpha$ chosen for the molecule were the C-C and the C-Cl vectors. Figure 4.8 shows the first Legendre polynomial for the C-C vector for the liquid and plastic phases. They both decay to zero, showing the orientational disorder, although the decay is faster in the liquid phase.

The RCF was also computed for the second Legendre polynomial and the reorientational times obtained from the integral of the RCF were computed.

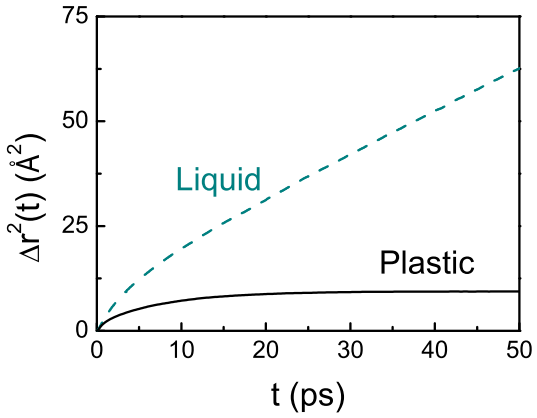

Figure 4.7: Mean-square displacement for the plastic (thick line) and liquid (dashed line) temperatures.

In the limit of rotational diffusivity, the ratio between τ_1 and τ_2 is equals to 3. It was found for the plastic phase: $\tau_1/\tau_2 = 2.55$ for the C-C vector and $\tau_1/\tau_2 = 2.28$ for the C-Cl vector. In the liquid phase: $\tau_1/\tau_2 = 2.56$ for the C-C vector and $\tau_1/\tau_2 = 2.58$ for the C-Cl vector. That shows that there are some jumps between orientations, and that in the liquid the reorientation is more diffusive.

In order to compare the short range ordering in the liquid and plastic phases of hexachloroethane the structure of the liquid is going to be detailed first and the plastic ordering will be compared next. The radial distribution function $g_{CM}(r)$ of the center of masses and the running coordination number $n_{CM}(r)$ were computed. Figure 4.9 presents $g_{CM}(r)$ and $n_{CM}(r)$ as a function of the distance r. The first two peaks of g_{CM} are shown. The classical definition of the first coordination shell (FCS) is the value of n_{CM} at the first minimum of g_{CM} (9.225Å): FCS = 13 molecules. This is a number that can be related to the closed packing on liquid structures (2). In order to study the local ordering, only the first neighbours were studied: the first maximum of g_{CM} (6.525Å) corresponds to $n_{CM} = 3.2$ molecules. Therefore, the analysis will be focused on the first 4 neighbours.

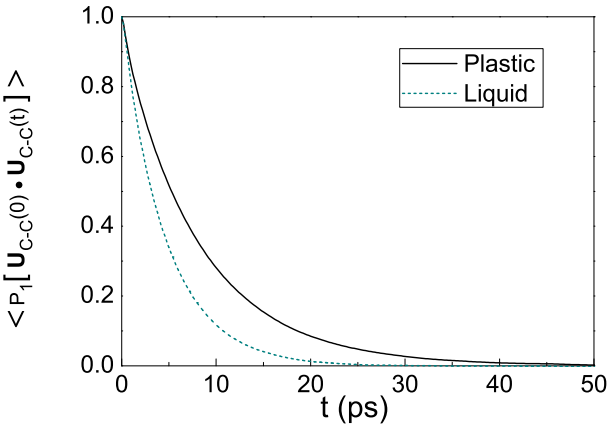

Figure 4.8: Reorientational correlation function for the C-C vector: plastic (thick line) and liquid (dashed line) temperatures.

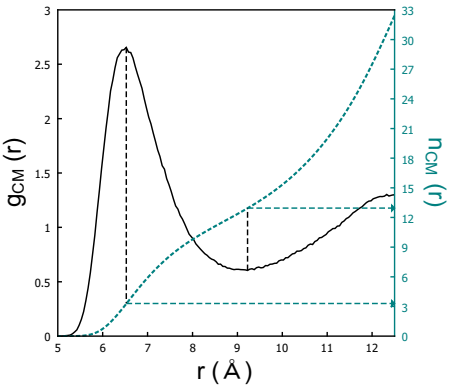

Figure 4.9: Short range order for the center of mass (CM) in the liquid C_2Cl_6. The radial distribution function g_{CM} (thick line) and the running coordination number n_{CM} (dashed line) are plotted as a function of the distance. The arrow at the first peak of g_{CM} marks n_{CM}=3.2 molecules and the first coordination shell (FCS) arrow shows that n_{CM}=13 molecules.

4.5.1 Positional ordering

The positional ordering distribution(θ_{pos}, ϕ_{pos}) of the first neighbour molecule is given in the left region of figure 4.10, with several local maxima, origi-

nated by two distinct main dispositions. First, from all the possible dispositions for the first neighbours, the most probable are in the equatorial region ($-0.8 < \cos(\theta_{\text{pos}}) < 0.8$) and some other in the polar (above and below of the reference molecule) region ($0.8 < |\cos(\theta_{\text{pos}})| < 1$). Having established this, from now on, both kind of positions have been considered, namely equatorial molecules and polar molecules. The equatorial molecules are arranged most likely around 72 and 107 degrees with respect to θ_{pos}. For ϕ_{pos} it is interesting to note that it does not prefer any location but some defined positions (0, $\pm$60, $\pm$120, $\pm$180 degrees). This analysis was extended for the first four neighbours obtaining the same result as indicated for the first neighbour.

4.5.2 Orientational ordering

To analyze the possible orientations two spots from the bivariate (θ_{pos}, ϕ_{pos}) positional map were chosen: one for the equatorial molecules and other for the polar molecules. Although only the result for one spot in each case is shown, the other spots contain the same information because of symmetry. For instance, molecules at a $\theta_{\text{pos}} = 72$ degrees have the same orientations as those at $\theta_{\text{pos}} = 107$ degrees. The same applies for the polar molecules.

In Figure 4.10 the bivariate distributions (θ_{ori}, ϕ_{ori}) for the different spots are shown. For the equatorial molecules there are two defined spots at $\cos(\theta_{\text{ori}}) = \pm 0.3$, ($\phi_{\text{ori}} = 120, -60°$) and some continuous probable values highlighted with a circle and a rectangle in the figure respectively. The last orientation degree of freedom, Ψ_{Ori}, completes the knowledge of the orientation. The map for the single spot is shown in the left side labeled as **(a)** along with one of the most probable orientations: all the values of θ_{ori} appear with probability $\neq 0$ although the preferred ones are centered around 0 and 180 degrees, while there are discrete orientations for Ψ_{ori} that are ($\pm$30, $\pm$90, $\pm$150 degrees). It is interesting to note the jumps of 60 degrees between probable orientations.

For the equatorial molecules with a continuous of possible θ_{ori} values, the result is shown and labeled as **(a)**. Unlike the first region here only

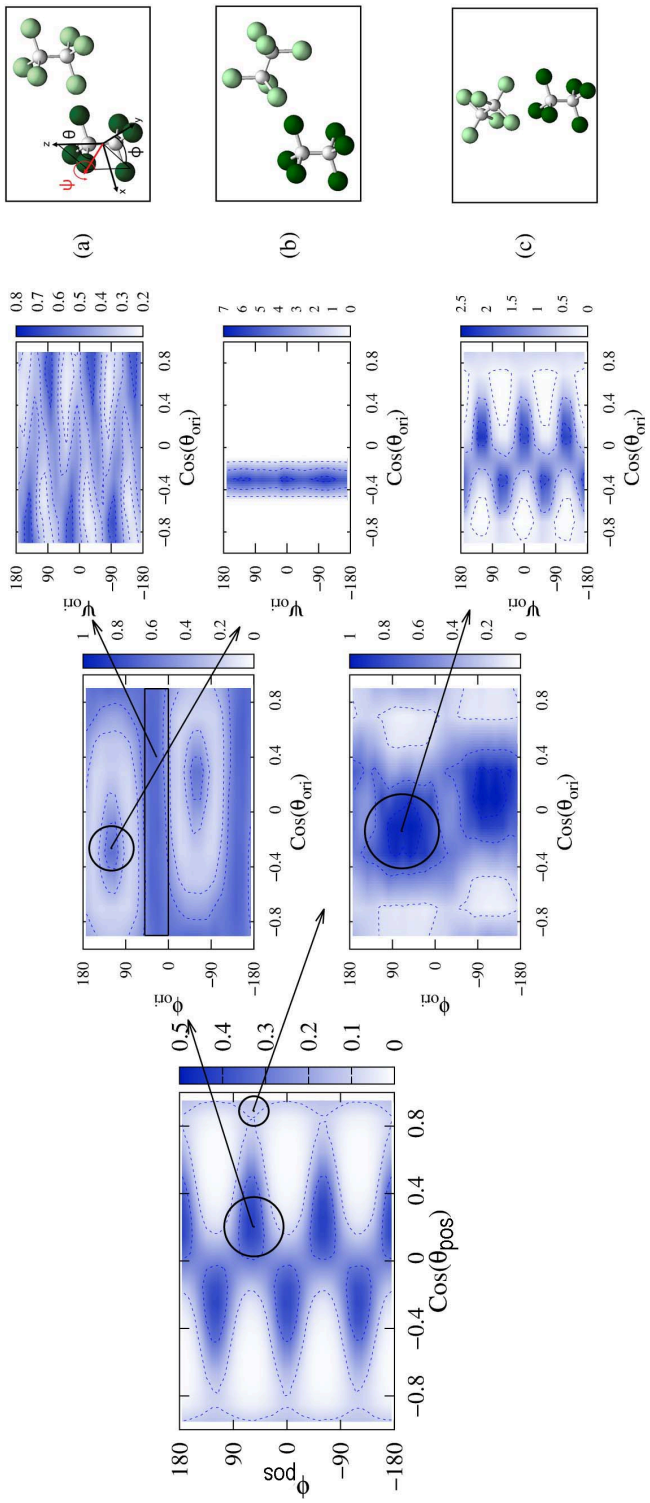

Figure 4.10: Bivariate maps for the Euler-Angles of the first neighbour. Left: Bivariate map for the position of the center of masses (CM) (θ_{pos}, ϕ_{pos}). Middle: Up the orientational bivariate map (θ_{ori}, ϕ_{ori}) for the equatorial molecules marked at Left. Down the orientational bivariate map (θ_{ori}, ϕ_{ori}) for the polar molecules marked at Left. Right: (a). Bivariate map (θ_{ori}, Ψ_{ori}) for the EQM labeled as $ori_{\parallel}$ and one of the most probable molecule orientation. (b). Bivariate map (θ_{ori}, Ψ_{ori}) for the equatorial molecules labeled as $ori_{\perp}$ and one of the most probable molecule orientation. (c). Bivariate map (θ_{ori}, Ψ_{ori}) for the polar molecules labeled as $ori_{\perp}$ and one of the most probable molecule orientation.

a narrow range of θ_{ori} values is relevant with a maximum at 107 degrees. Again Ψ_{ori} shows some discrete values (0 and $\pm$120 degrees). Recalling that ϕ_{ori} and Ψ_{ori} contain the relative orientations of chlorides, this can be understood as a fact of the molecular symmetry: chlorides avoid facing other chlorides, appearing most probably every 60 degrees.

The analysis of the polar molecules shows two big areas centered about $\cos(\theta_{\text{ori}}) = \pm 0.2$ and $\phi_{\text{ori}} = 60$ and -120 degrees. The bivariate distribution (θ_{ori},Ψ_{ori}) labeled as **(c)** reveals an equivalent arrangement to that of **(b)**.

There are two different preferred orientations for the equatorial molecules labeled as **(a)** and **(b)** in Figure 4.10. **(a)** shows the chloride shell pointing in the same direction and will be labeled them as ($ori_{\parallel}$), while **(b)** shows the chloride shell pointing in perpendicular directions and will be labeled it as ($ori_{\perp}$). In the case of the polar molecules, **(c)**, the carbon bone is oriented perpendicular as **(b)**. The orientational analysis for the first four neighbours was repeated, finding the same results as those for the first neighbour. This points out that the C_2Cl_6 has a complex rather than simple translational and orientational structure and brings out some questions about to what extent the short range order of the plastic phase is persistent.

In order to compare the probability distributions of the $ori_{\parallel}$ and $ori_{\perp}$ structures, two questions are investigated: are they equiprobable? and do they appear at the same distances?. To answer these questions, the population (percentages) in each configuration and the characteristic distance for each neighbour were analyzed. As usual, the first neighbour was studied. The distribution of distances for the CM in the $ori_{\parallel}$ and $ori_{\perp}$ classes was computed. The result is a gaussian distribution plotted in Figure 4.11a. The dotted and solid lines correspond to $ori_{\parallel}$ and $ori_{\perp}$ molecules respectively. The second question is already answered here, the two structures appear at different distances. The most probable values are $d_{\parallel} = 5.92$Å and $d_{\perp} = 6.045$Å. Regarding to the population of each class it was found that 52% of the molecules are oriented as the class $ori_{\parallel}$ and 14% are $ori_{\perp}$, this answers the first question: the parallel orientation is more probable in the

first neighbour. This concludes that they are non-equivalent structures.

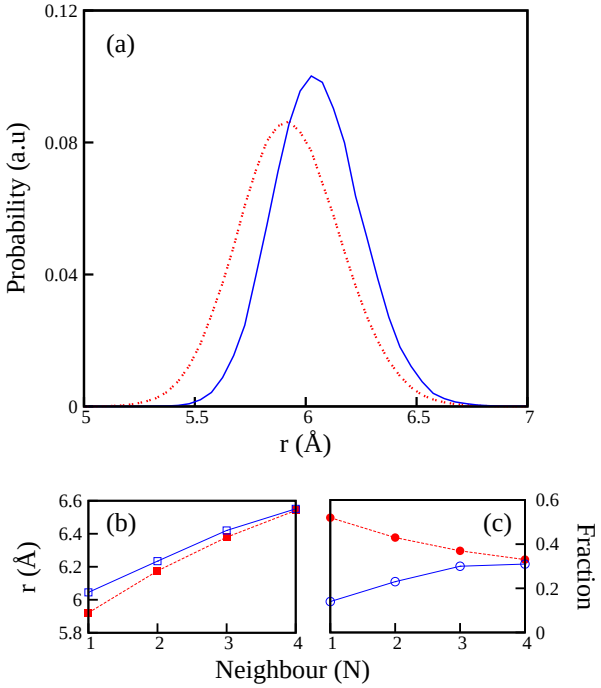

Figure 4.11: Differences of the $ori_{\parallel}$ (dotted line) and $ori_{\perp}$ (solid line) classes. (a). Distance distribution of the center of mass CM for the first neighbour. (b). Most probable value of distance of CM for the $ori_{\parallel}$ and $ori_{\perp}$ classes (filled and empty squares respectively) as a function of the number of neighbours. (c). Fraction of molecules populating each $ori_{\parallel}$ and $ori_{\perp}$ class (filled and empty circles respectively) as a function of the number of neighbours.

The results for the first four neighbours are summed up in figures 4.11b and 4.11c. Panel b shows the distance at which the close neighbour appears for the $ori_{\parallel}$ and $ori_{\perp}$ classes (dotted and solid lines respectively). The behaviour of the first neighbour is repeated here but the distances get closer as n_{CM} increases. Regarding to the population of each class, the values are plotted in Figure 4.11c. A similar tendency is seen, both classes approach to the same population as going from the first to the fourth neighbour. The amount of molecules in a given orientation is dependent on the distance, which can be related to the minimal configuration energy. Also, the non-equivalency of both structures is strong in the first neighbours and dis-

appears as n_{CM} grows. Finally for the fourth neighbour the common pattern in liquids is recovered where different arrangements are equiprobable and they are at the same distance.

4.5.3 Short range order comparison to the plastic phase

The Euler angles analysis done for the liquid phase of C_2Cl_6 was repeated as in the previous section for the first neighbour in the bcc plastic phase. The results are shown in Figure 4.12.

From the results in Figure 4.12 it is shown that the first neighbour position and orientation is really similar but for some details. The positions of the bcc are better defined and correspond to the expected unit cell (see the inset in the center of Figure 4.12). A remarkable result is that the orientational order in the liquid is the same as that of the bcc phase. The same most probable orientations are present in both phases. The difference lies in the fact that molecules in the liquid can interchange continuously between some of the orientations, while in the plastic phase there are some allowed orientations and the molecules undergo "jumps" between orientations.

The orientational ordering for the first eiqht neighbours in the plastic phase was studied: the results for neigbours 1-8 are the same. This is expected since they are the nearest neighbours in the diagonals of the cubic lattice. The representation of the bcc lattice is shown as an inset in the center of Figure 4.12, it was generated by taking the possible orientations as for the liquid structure, but for the eight neighbours. The possible orientations are, as in the liquid study, parallel and perpendicular orientations. Remember from last subsection, that it was found some molecules in the polar region of the maps, being above the C-C bonds. In the lattice bcc phase, where the reference molecule is oriented in the [111] diagonal, the polar molecules are the ones in the vertices of that [111] diagonal. It was also found that the Cl atoms lie in the [100] direction. These are the same results that Criado & Muñoz found for the bcc structure (1).

The possible orientations in the bcc phase appear every 60 degrees. This

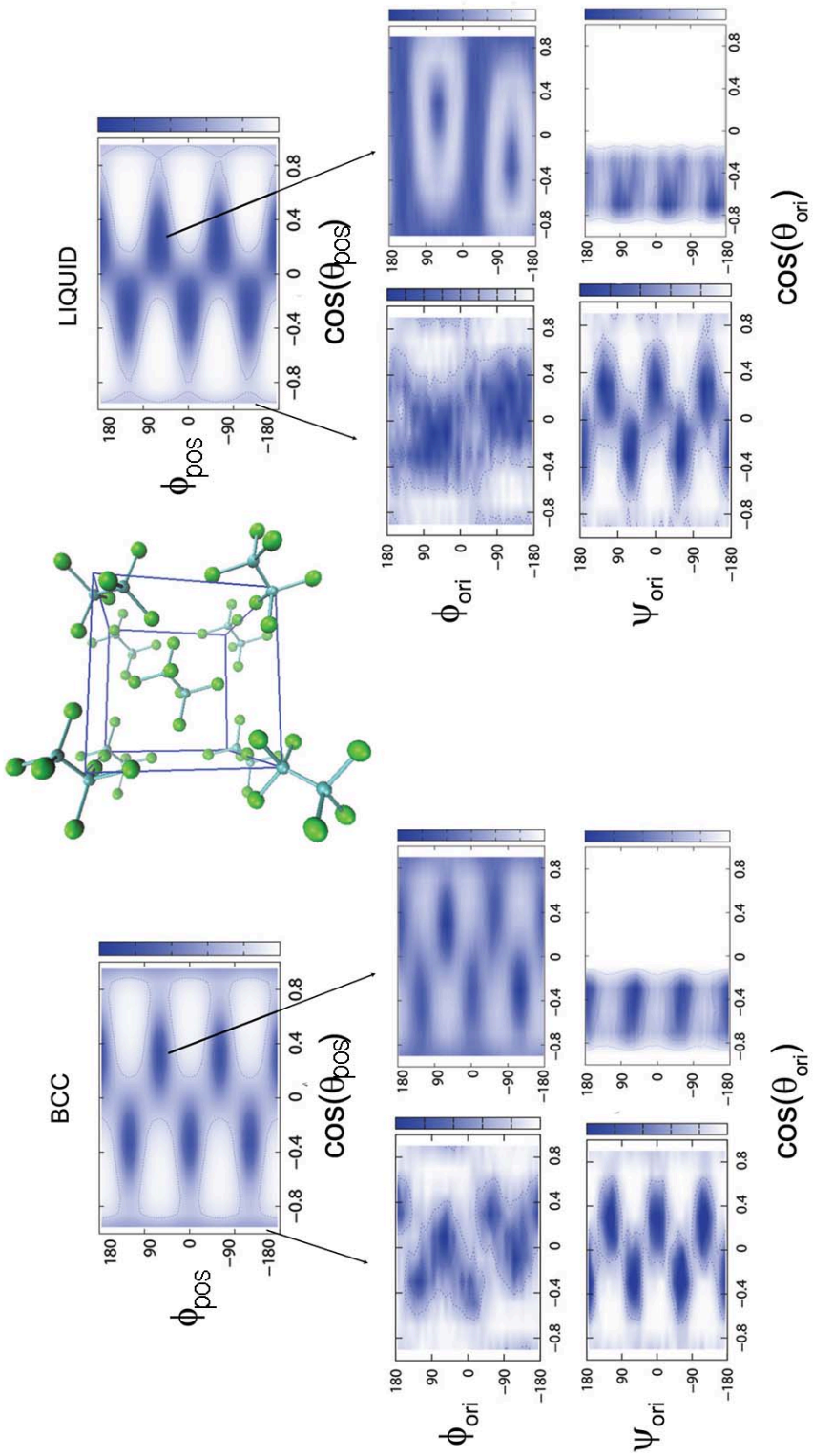

Figure 4.12: First neighbour analysis for the plastic bcc (left) and liquid (right) phases of C_2Cl_6. At the top, the positional bivariate maps ($cos\theta_{pos}$, ϕ_{pos}). In the middle, the orientational ($cos\theta_{ori}$, ϕ_{ori}) maps. At the bottom, the orientational ($cos\theta_{ori}$, ψ_{ori}) maps. The arrows show the regions that are analyzed: one of the equatorial positions and one of the polars, as it was studied for the liquid in the last section. The inset of the figure in the center shows the corresponding bcc unit cell, with some of the possible orientations.

agrees with the angle between chlorides atoms. The molecules avoid facing their chlorides because of the repulsion that is mainly steric. The next step to go further in the analysis is to compare the long range order for both phases, the results are shown in the next subsection.

4.5.4 Long range order

The tools of information theory introduced in chapter 3 were implemented to obtain a comparison in the long range order, i.e for neighbours beyond the first coordination shells. This is a difficult task to tackle using the detailed analysis done for the first neighbour. This is why information theory is a valuable source of analysis. The investigation was limited to the positional order, using the positional maps obtained from Euler angles up to the neighbour 75. The results are shown in Figure 4.13.

Figure 4.13a) shows:

- At the top, the radial distribution function (RDF) against the number of neighbours (MCN). The use of MCN instead of the distance is done to remark the presence of different coordination shells. This is done for both, liquid and bcc phases.

- Entropy (H) associated to the positional maps, see equation 3.1 in chapter 3. This, for each phase.

- Mutual Information (I_2) according to equation 3.4. In this case measuring the correlation between the positional θ_{pos} and ϕ_{pos} angles in each phase.

- At the bottom, an additional measurement is shown:

  The Kullback-Leibler divergence (KL) to measure the "resemblance" between the liquid and the bcc phases:

$$D_{KL} = \sum p(x) log \frac{p(x)}{q(x)} \tag{4.12}$$

  where q(x) will be the probability distribution associated to the liquid phase being compared against the more ordered bcc (p(x)).

Figure 4.13b) shows six positional maps (for the bcc and liquid phases) corresponding to different MCN that contain interesting information:

1. First neighbouring shell limit in bcc (Neighbour 8th).
2. First neighbouring shell limit in bcc (Neighbour 14th).
3. Neighbour 26th because of its low KL distance.
4. Neighbour 27th because a low KL distance and the same I_2 value.
5. Neighbour 43th because of a low KL distance but different I_2 values.
6. Neighbour 60th because of its high KL distance.

Next, the results for each of the quantities computed and explained above are detailed. The results are shown separately for the sake of a deeper description and a better understanding.

Radial distribution function RDF: The RDF functions shown in Figure 4.13 present information on the positional order of the center of mass of neighbouring molecules. In the case of the bcc plastic phase (red thick line), it shows well defined shells, although not all of them have an RDF that goes to zero as it would be the case for a crystal. The RDF shows the positional order that is preserved for long distances and at the same time it shows a broadening of the peaks as a result of the orientational disorder and thermal vibrations for the position. The first two neighbouring shells appear at 8 and 14 molecules, the nearest, and next-nearest neighbours in the bcc lattice. In the case of the liquid (blue dashed line), the RDF shows less structure. It has a broader first peak and it decays in the long range order compared to the bcc. However, at a MCN of 14 it mimics the same shell definition (a low value of RDF) as the bcc phase.

Shannon Entropy H***:*** The Shannon entropy H for the positional maps (variables $cos(\theta_{pos})$ and ϕ_{pos}) rescaled by its maximum value H_∞. Remember from equation 3.1 that H is maximum and negative for a PDF with no information (a disordered structure in our case) and zero for a PDF with maximum information (ordered system). In the case of the rescaling

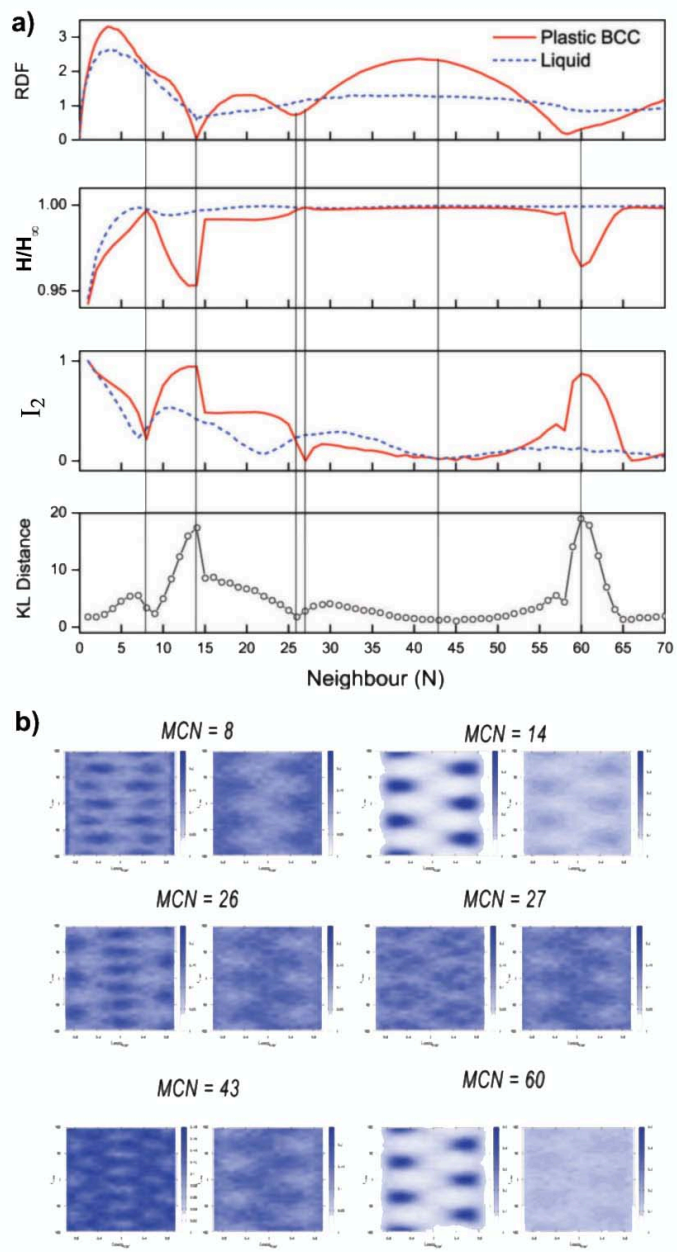

Figure 4.13: Information theory analyses to compare the positional maps ($cos\theta_{pos}$, ϕ_{pos}) between the bcc (red thick line) and liquid (blue dashed line) phases. a) From top to bottom: Radial distribution function (RDF), the entropy rescaled to the maximum value (H/H_∞), the mutual information (I_2) and the Kullback-Leibler divergence (KL). All the former quantities are represented as a function of the mean coordination number (MCN). b) Bivariate positional maps ($cos\theta_{pos}$, ϕ_{pos}) at 6 different MCN values, shown with lines in a). For each MCN the maps both phases are compared, the bcc plastic phase in the left and the liquid phase in the right.

H/H_∞, the function goes to 1 for a completed disordered system (gas like) and lower values indicate an increasing positional order. An interesting feature for the case of the bcc plastic phase is the periodic appearance of low values of H/H_∞ that show the periodic ordering of the lattice. For instance, the values are similar between the 1st and 3rd shells and between the 2nd and 4th shells. In the case of the liquid phase it is notorious a higher ordering for the first neighbours. An increase in the ordering around the 12th neighbour can also be seen, again it mimics the plastic phase as in the RDF. But in the long range the liquid looses the positional ordering.

Mutual information I_2***:*** The value I_2 quantifies the correlation between the positional variables $cos(\theta_{pos})$ and ϕ_{pos} (chapter 3, equation 3.4). Going from 1 to 0 for the case of a full correlation to no correlation for zero. The first aspect to note is that the I_2 is correlated to the entropy H: Low values of H/H_∞ indicate higher order, meaning a higher correlation between the positional variables. The value I_2 is more structured for the positional order in the liquid compared to H (It does not simply decay). For instance, at an MCN of 14 the positional maps shown in Figure 4.13b are similar between both phases, although the proportion of molecules that poblate that state is lower in the liquid phase. The periodic ordering of the bcc lattice can also be seen when comparing the positional maps of the neighbours 14th and 60th. One interesting thing from these results is that I_2 shows clearer the ordering inferred from the positional maps than the Shannon entropy.

Kullback-Leibler divergence KL***:*** The KL quantity can measure the resemblance between the positional maps of liquid and bcc phases. A lower value of KL indicates a higher similarity. The two maxima of this function appear for the neighbours 14th and 60th, for which the positional ordering is better defined in the bcc phase. This agrees with the other quantities at the same MCN values, a low RDF value, low entropy and high mutual information. Although from the maps there is still some resemblance of the liquid phase to the bcc, the peaks are lower, which means that most of the molecules do not show that ordering. Also, the low KL values for the first

neighbours confirm the similarity in the positional order for both phases as it was shown in the previous subsection.

Finally, the above results show that theory information presents valuable information without going into detail on the Euler angles analysis.

4.6 Summary and conclusions

The structure of the C_2Cl_6 liquid and bcc plastic phases was studied defining unambiguously the operations over translational and rotational states using the proper Euler-angles. The gromos53a6 (17) force field used in this work shows better results than previous works, when compared to experimental results on the structure of the plastic phase.

The detailed analysis on the first neighbour showed that the positional and the orientational ordering is really similar to that of the plastic phase. For this molecule, for which the steric effects are the most important it shows that in the short range order the liquid is ordered. The molecules try to have a close packing and the possible orientational configurations are leaded by the shape of the molecules: avoiding facing chloride shells.

Information theory is a valuable and important tool to analyse the results obtained with the Euler angles study. From this kind of analysis, it confirms the results that otherwise should have been obtained by a close inspection of the maps. This would be a demanding task. It is also important, that the analyses can also be quantitative and not merely qualitative. When analyzing the information about the positional maps, the mutual information showed a better insight on the structure that the Shannon entropy.

The use of the Euler angles and information theory analyses lead to a deeper and new understanding on the structure hidden on disordered phases like the liquid and plastic crystal studied here.

4.7 Bibliography

[1] A. Criado and A. Muñoz. A molecular dynamics simulation of the plastic phase of hexachloroethane. *Molecular Physics*, 83(4):815–833, 1994. 54, 62, 63, 72

[2] J. Finney. Random packings and the structure of simple liquids. i. the geometry of random close packing. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 319, pages 479–493. The Royal Society, 1970. 66

[3] H. E. Fischer, G. J. Cuello, P. Palleau, D. Feltin, A. C. Barnes, Y. S. Badyal, and J. M. Simonson. *Appl. Phys. A: Mater. Sci. Process.*, 74: S160, 2002. 57

[4] H. E. Fischer, A. C. Barnes, and P. S. Salmon. Neutron and x-ray diffraction studies of liquids and glasses. *Reports on Progress in Physics*, 69(1):233, 2005. 59

[5] P. Gerlach and W. Prandl. Short-range-order correlations in the orientationally disordered phase of hexachloroethane. i. diffuse x-ray scattering. *Acta Cryst.*, A44:128–135, 1988. 54

[6] P. Gerlach, D. Hohlwein, W. Prandl, and F. W. Schulz. The plastic phase of hexachloroethane, C_2Cl_6: A neutron powder and single-crystal investigation. *Acta Cryst.*, A37:904–908, 1981. 54, 62, 63

[7] P. Gerlach, W. Prandl, and J. Lefebvre. The plastic state of C_2Cl_6. a comparison of a monte carlo simulation of the molecular distribution with single crystal neutron data. *Molecular Physics*, 49(4):991–999, 1983. 54

[8] P. Gerlach, B. Dorner, W. Prandl, and J. Lefebvre. Short-range-order correlations in the orientationally disordered phase of hexachloroethane. ii. elastic and quasielastic neutron scattering. *Acta Cryst.*, A44:251–257, 1988. 54

[9] H. Goldstein, C. P. Poole, and J. L. Safko. *Classical Mechanics*. Addison Wesley, San Francisco, third edition edition, 2002. 56

[10] A. Henao, M. Rovira-Esteva, A. Vispa, J. L. Tamarit, E. Guàrdia, and L. C. Pardo. Insights into the determination of molecular structure from diffraction data using a bayesian algorithm. *J. Phys.: Condens. Matter*, 25:454217, 2013. 53

[11] A. Henao, S. Pothoczki, M. Canales, E. Guàrdia, and L. C. Pardo. Competing structures within the first shell of liquid C_2Cl_6: A molecular dynamics study. *J. Mol. Liq.*, 190:121–125, 2014. 53

[12] B. Hess, C. Kutzner, D. V. der Spoel, and E. Lindahl. Gromacs4: Algorithms for highly efficient, load-balanced and scalable molecular simulation. *J. Chem. Theory Comput.*, 4:435–447, 2008. 55

[13] M. Howe, R. McGreevy, and P. Zetterström. Computer code correct. *Correction program for neutron diffraction data, NFL Studsvik*, 1996. 57

[14] Y. Morino and M. Iwasaki. The estimation of the hindering potential barrier of hexachloroethane by electron diffraction investigation. *J. Chem. Phys.*, 17:216, 1949. 54, 56

[15] A. Muñoz and A. Criado. A molecular dynamics interpretation of the inelastic neutron spectra in the plastic phase of hexachloroethane. *Molecular Physics*, 84(6):1207–1225, 1995. 54

[16] P. Negrier, J. L. Tamarit, M. Barrio, and D. Mondieig. Polymorphism in halogen-ethane derivatives: Ccl3–ccl3 and clf2c–cf2cl. *Crystal Growth & Design*, 13(2):782–791, 2013. 54, 62

[17] C. Oostenbrik, A. Villa, A. E. Mark, and W. F. V. Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: The gromos force-field parameters sets 53a5 and 53a6. *J. Comp. Chem.*, 25:1656–1676, 2004. 55, 78

[18] L. C. Pardo, J. L. Tamarit, N. Veglio, F. J. Bermejo, and G. J. Cuello. Comparison of short-range-order in liquid-and rotator-phase states of a simple molecular liquid: A reverse monte carlo and molecular dynamics analysis of neutron diffraction data. *Physical Review B*, 76(13):134203, 2007. 57

[19] G. Placzek. The scattering of neutrons by systems of heavy nuclei. *Physical review*, 86(3):377, 1952. 57

[20] M. Rovira-Esteva, N. Murugan, L. Pardo, S. Busch, J. L. Tamarit, S. Pothoczki, G. Cuello, and F. J. Bermejo. Interplay between intramolecular and intermolecular structures of 1, 1, 2, 2-tetrachloro-1, 2-difluoroethane. *Physical Review B*, 84(6):064202, 2011. 59, 60, 64

[21] P. Rozier, A. Burian, and G. Cuello. Neutron and x-ray scattering studies of li 2 o–teo 2–v 2 o 5 glasses. *Journal of non-crystalline solids*, 351(8):632–639, 2005. 59

[22] S. Seki and B. Momotani. Heats of transition of hexachloroethane. *Bulletin of the Chemical Society of Japan*, 23:30–31, 1950. 55

[23] S. Seki and M. Momotani. Heats of transition of hexachloroethane. *Bulletin of the Chemical Society of Japan*, 23(1):30–31, 1950. 54

[24] Y. L. Slovokhotov, A. S. Batsanov, and J. A. Howard. Molecular van der waals symmetry affecting bulk properties of condensed phases: melting and boiling points. *Struct Chem*, 18:477–491, 2007. 55

[25] D. A. Swick, I. L. Karle, and J. Karle. The structure and internal motion of hexachloroethane. *J. Chem. Phys.*, 22(7):1242–1245, 1952. 54, 56, 60, 61, 62

Chapter 5

Short range order in liquid water

5.1 Motivation

The local structure of liquid water is still a controversial subject. In contradiction to the classical picture of a symmetric tetrahedral local order of neighboring water molecules around a central one (10, 11, 16, 17), two main alternative descriptions have been proposed: a local order with only one acceptor and one donor that would lead to a chain and ring structure of liquid water (7) and a two-state model that would imply an inhomogeneous picture of the water structure (6, 8, 13). Kühne *et al.* proposed recently (11) a way to reconcile the classical picture of the short range order of water with its counterparts by means of an asymmetry of the electronic interaction between molecular contacts. This asymmetry was found in the interaction between the two strongest hydrogen bond donors and acceptors. Recently Kühne *et al.* have also found a difference in the geometry of these hydrogen bonds (12). In this chapter, a careful analysis of molecular dynamics simulations is performed to show that, beyond the differences on H-bond geometries, there is also a difference in the relative orientation of two water molecules that go together with a difference in the contact energy.

5.2 Computational details

5.2.1 simulation details

Molecular Dynamics (MD) simulations were performed using a system of 512 water molecules at T=298 K and ρ=0.9979 g/cm^3 using the TIP4P/2005 water model (1). This is a rigid non-polarizable model which reproduces many properties of liquid water and ice (14, 19). The same analysis was also performed on simulations of the widely used SPC/E water model (2) and the flexible TIP4P/2005f water model (9). Long-range electrostatic interactions were calculated with the Ewald summation technique. The time step was set to 2 fs in the case of the rigid water models and 0.1 fs in the case of the flexible water model. The simulations consisted of an equilibration run of 100 ps followed by a production run of 1000 ps. During the production run, the configurations were stored every 500 time steps for further analysis. The three models: TIP4P/2005, its flexible version TIP4P/2005f and the SPC/E model have been used to do the analysis of the molecular arrangement of liquid water. The results given by the three are qualitatively the same, therefore only the results for the TIP4P/2005 model are shown.

5.2.2 Axes definition

The probability distribution function $g(r, \Omega_{\text{pos}}, \Omega_{\text{ori}})$ were obtained by studying the distance between oxygens of two water molecules and defining an axis set to each molecule. In order to chose the coordinate system to the molecule it is important to note that the Euler angles ϕ_{ori} and ψ_{ori} become degenerate at the poles. In order to avoid regions of interest to be situated at the poles, two definitions of orthonormal axes sets were used in this study: To study the relative *position* of two molecules, the z axis was defined along the dipole of the water molecule, x perpendicular to the H-O-H plane and y in the H-O-H plane (see inset of figure 5.1). As for the relative *orientation*, however, a different set of coordinate axes was chosen with z perpendicular to the H-O-H plane, x parallel to the H-O-H plane and y along the molecular dipole.

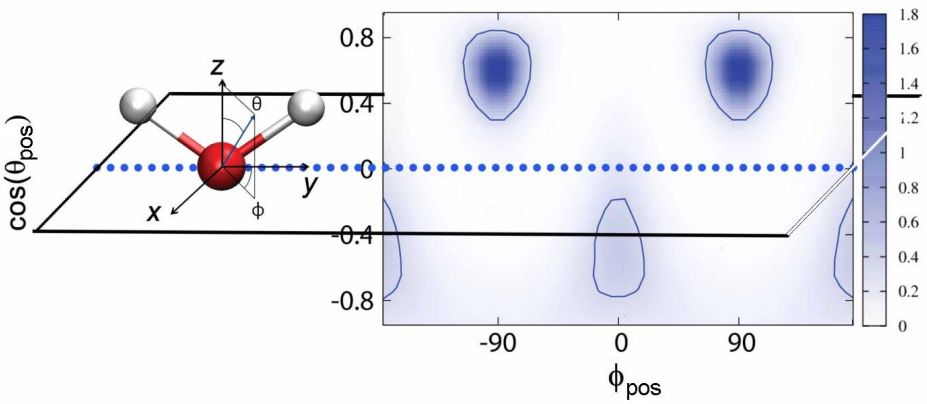

Figure 5.1: Position distribution function $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}})$ describing the probability of finding a molecule in a certain region of space in spherical coordinates. The contour for $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}}) = 0.3$ is shown. The axes chosen to study the position are shown in the inset of the figure.

5.3 Positional ordering in the first hydration shell

The first quantity that describes the short range order of the molecules is their positional arrangement around a central molecule. Molecules in a liquid are changing continuously their positions; however, there are some preferred locations. These preferred locations can be visualized with the distribution function $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}})$ which is shown for the first four neighbors surrounding a central molecule in figure 5.1. Hydrogen bond acceptors are located in the northern hemisphere ($\cos(\theta_{\text{pos}}) > 0$) and donors in the southern hemisphere ($\cos(\theta_{\text{pos}}) < 0$). For the acceptors, the highest probabilities to find a neighbouring molecule are at $\theta_{\text{pos}} = 52.28°$ which is half the H-O-H angle of the TIP4P/2005 water molecule and very close to the tetrahedral angle $\theta_t/2 = 54.7°$. On the other side, donors have a broader distribution than acceptors that indicates a higher positional disorder. Previous studies have also found the donor disorder using a variety of water models and simulation methods (4, 15, 18). This might be related to the capability of classical force-fields to encode the different distribution of acceptor and donor electrons, the so-called negativity track between the lone pairs of a water molecule (12). It is possible to quantify this difference

between donor and acceptor distributions using the Kullback-Leibler divergence from information theory which is shown next. The same quantity was used in chapter 4 to quantify the differences between the liquid and plastic phases of Hexachloroethane.

5.3.1 Quantifying the donor/acceptor distribution asymmetry

Figure 5.1 shows that there is a different distribution of oxygen atoms belonging to donors and acceptors around a central water molecule. This distribution is encoded in the 2D probability distribution function $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}})$, where the acceptors belong to the north hemisphere and donors belong to the southern hemisphere. To quantify the difference between the distribution of donors and acceptors means to calculate how different are the distributions of northern and southern hemispheres. Information theory provides with a tool to calculate the "distance" between two different probability distributions: the Kullback-Leibler divergence (D_{KL}) defined as:

$$D_{KL}\left(P||Q\right) = \sum_i P_i \ln \frac{P_i}{Q_i} \tag{5.1}$$

Although D_{KL} is not a real distance because it is not symmetric, i.e. $D_{KL}\left(P||Q\right) \neq D_{KL}\left(Q||P\right)$, it is possible to define a symmetrized version as $D_{KL}(P,Q) = 1/2\left(D_{KL}\left(P||Q\right) + D_{KL}\left(Q||P\right)\right)$. In the present case the three distances have been checked (the two non-symmetrized and the symmetrized one) obtaining approximately the same results and exactly the same trends. In order to be able to calculate this distance an alternative axis set to the one of figure 5.1 have been chosen, so that donors are located in the position of the acceptors (see inset of figure 5.2). Figure 5.2 shows the results for the symmetrized D_{KL} as a function of the Molecular Coordination Number (MCN), i.e. the number that results after ordering molecules by increasing distances. As expected, the first four neighbors have the maximum possible asymmetry in the liquid. However, after the fifth neighbour the asymmetry increases again having a maximum at $MCN = 8$ and finally fades out for $MCN > 20$. To out that this asymmetry is a trivial effect due to an increasingly loose of correlation with distance, D_{KL} have been

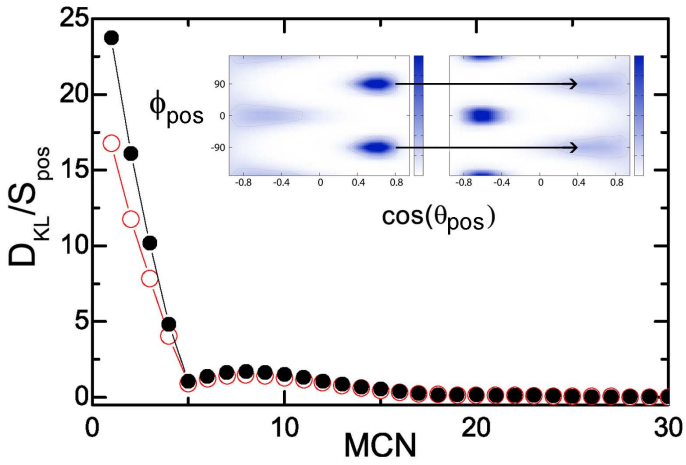

Figure 5.2: The symmetrized Kullback-Leibler divergence between the distribution of donors and acceptors as a function of the Molecular Coordination Number. The inset shows the maps where alternatively donors and acceptors occupy the same spots of the 2D probability distribution $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}})$.

divided by the entropy associated to the maps ($S = \sum_i P_i \ln P_i$) obtaining again the same trend.

5.4 Orientational ordering in the first hydration shell

The orientation of neighbouring molecules is encoded in the three-dimensional function $g(\cos(\theta_{\text{ori}}), \phi_{\text{ori}}, \psi_{\text{ori}})$ which shows the distribution of orientations of the neighbouring molecules at a given position. The acceptors and donors orientations are evaluated separately. For the following analysis, two groups were made: one contains the two closest neighbouring donors and the other the two closest acceptors.

Both donor and acceptor orientational distributions can be visualized as 3D isocontour surfaces. The isocontour surface for hydrogen bond acceptors is shown in figure 5.3, the one for donors is virtually the same. In both

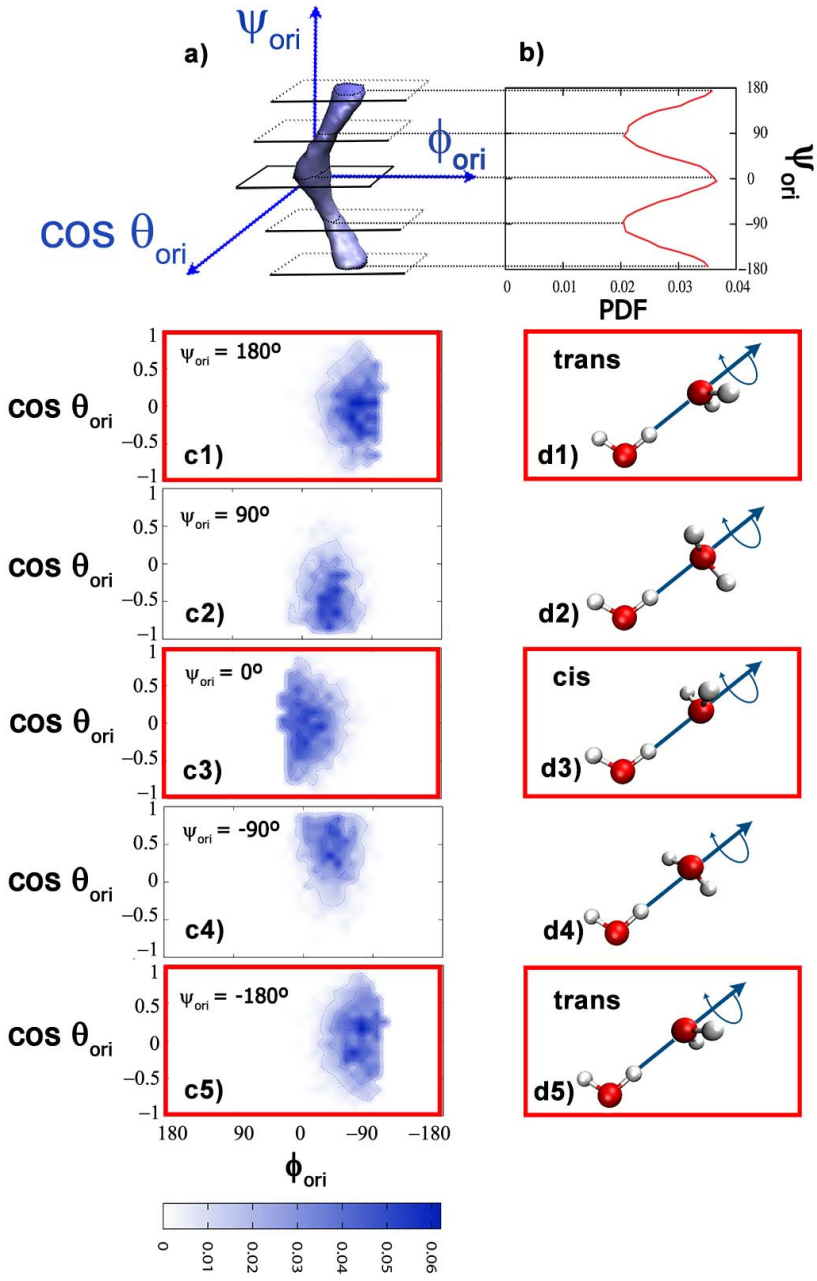

Figure 5.3: (a) Three dimensional probability distribution $g\left(\cos(\theta_{\rm ori}), \phi_{\rm ori}, \psi_{\rm ori}\right)$ describing the orientation of a water acceptor molecule. (b) Probability of successive cuts of $g\left(\cos(\theta_{\rm ori}), \phi_{\rm ori}, \psi_{\rm ori}\right)$ along planes perpendicular to the $\psi_{\rm ori}$ axis. Panel (c) shows $g\left(\cos(\theta_{\rm ori}), \phi_{\rm ori}\right)$ for successive cuts of the probability distribution of panel (a) at different values of $\psi_{\rm ori}$. Panel(d) shows the orientations obtained from the maximum of $g\left(\cos(\theta_{\rm ori}), \phi_{\rm ori}\right)$.

cases the surface resembles a spiral winding around the z axis. Inside the isocontour surface, there is a continuum of possible molecular orientations. The part of orientational space which is not inside the isocontour surface is explored less frequently by the molecules and there are even orientations that were never observed in the presented simulations. Since the spirals enclose a continuous volume, it is possible to transform each probable orientation into another one continuously along a path within the spiral, i. e. there is a certain sequence of orientations through which the molecules pass when they change the orientation. The question that will be studied in the following is: Which are the most probable orientations for a hydrogen bond acceptor (or donor) to be in?

To answer this question, i. e. to find the most probable orientation of the molecules, 2D slices were cut through the orientational distributions in figure 5.3 to obtain a view inside the isocontour surfaces. These slices were in both cases cuts along horizontal xy planes of the 3D distributions shown in figure 5.3a at several heights, i. e. several values of z. In the case of hydrogen bond acceptors, this corresponds to a series of $g(\cos(\theta_{\text{ori}}), \phi_{\text{ori}})$ at several ψ_{ori} values (see figure 5.3b). Exemplary cuts are shown below their corresponding 3D isocontour surfaces in figure 5.3b. As it can be seen in the figure, successive slices correspond to a rotation of the molecule around the O-H axis of the central molecule. In panel b the probability for each slice is shown, that is the probability to find a particular orientation. The orientations with maximum probability are highlighted in the same figure: $\psi_{\text{ori}} = \pm 180°$ and $\psi_{\text{ori}} = 0°$.

The two most likely orientations for both dimers – formed by a central molecule and either an acceptor or a donor – are therefore two possible configurations: one with a parallel dipole alignment and another one with an angle of about $\Phi_{\text{dip-dip}} \approx 110°$ between dipoles. In analogy to the classification of conformations in hexagonal ice (3), the parallel dipole alignment will be referred to as *cis* while the one with a $\Phi_{\text{dip-dip}} \approx 110°$ angle between them will be called *trans*, cf. figure 5.3d. The *trans* configuration is very close to the most stable water dimer in vacuum (5). Both donors and acceptors, were found with equal probability in a *cis* or *trans* configuration so that the *cis*:*trans* ratio was in either case 50:50 (in hexagonal ice, the

proportion of *trans* pairs is higher with a *cis*:*trans* ratio of about 40:60 (3)).

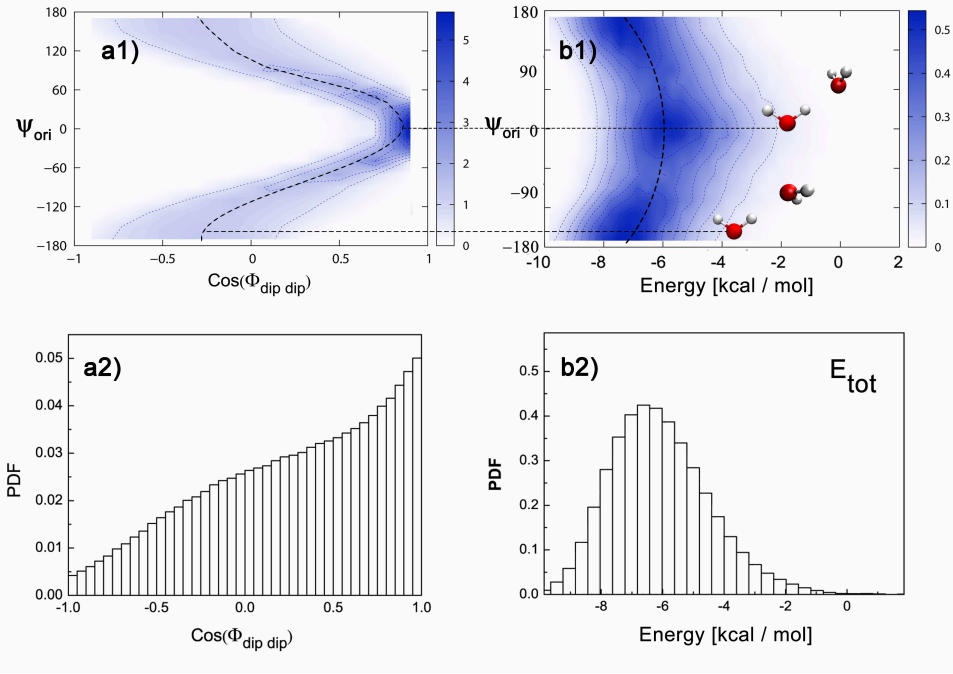

Figure 5.4: Two dimensional probability maps for the relative dipole orientation $\Phi_{\text{dip-dip}}$ (a1) and energy (b1) of an acceptor as a function of the Euler angle ψ_{ori}. Panels a2 and b2 show the projection of the data in figures a1 and b2 onto the x axis, i. e. the probability density functions of $\Phi_{\text{dip-dip}}$ and of the energy of a dimer involving a neighbouring acceptor molecule.

Both dimers were characterized by measuring the correlations between conformations and interaction energies as well as to their correlation with dipole moment orientations. It must be pointed out that the calculation of the energy between two water molecules is done using a pairwise additive potential. If a polarizable model were used, the effect of the surrounding molecules within the first hydration shell would explicitly be taken into account and would probably lead to a stronger correlation between the orientation of the deemed molecules. The 2D probability distribution functions of the energy and relative dipole orientation $\cos(\Phi_{\text{dip-dip}})$ of the dimer with the angle ψ_{ori} was calculated and is shown in figure 5.4. The 1D probability distributions of these quantities integrated over all angles ψ_{ori} are shown in the same figure (for one acceptor, similar results were obtained for the donors). From figure 5.4 it is clear that the relative dipole orientation of

two molecules can be understood as the combination of a broad peak associated with the *trans* conformation where dipoles are approximately in a perpendicular orientation together with a narrower peak associated to *cis* dimers where dipoles are parallel. Such a bimodal distribution is not visible on first sight in the 1D probability distribution of the energy. Only if plotted as a function of angle ψ_{ori} it becomes clear that the total energy for an acceptor consists of tightly linked dimers in *trans* conformation with almost perpendicular dipoles and more loosely linked dimers in *cis* conformation with almost parallel dipoles. The difference in energy between the *cis* and *trans* dimers calculated between the maxima of the two 2D distribution $P(E, \psi_{\text{ori}})$ for both dimers is 1.3 kcal/mol.

Kühne *et al.* also found an asymmetry between the contacts of water molecules (in that case the energy difference was about 2.5 kcal/mol) which was associated to an electronic criterion and recently to a different geometry of the hydrogen bond (12). In this chapter there was also found that the differences in energy have an electronic origin, they result from different dipole-dipole interactions between molecules which in turn have a clear geometric origin with the relative orientation of the two molecules. However it has to be considered that these two asymmetries might arise from different physical origins since our simulation is purely classical.

5.5 Correlations beyond the dimers

In this section, the question about if beyond the evident difference in molecular orientation, the existence of two different dimers causes any differences in the ordering of the molecules around them is addressed. This is done both within the first hydration shell and beyond.

The first question is whether the presence of any of the two different dimers has any effect on the position or orientation of the other molecules in the first hydration shell. To answer it, one acceptor (donor) involved in a *cis* or *trans* dimer was chosen and the Euler angles analysis was repeated for the other three molecules: no appreciable differences was found. This

means that if a neighbouring water molecule is in a certain configuration, it does not impose any restriction to the relative orientation of the other hydrogen bonded molecules.

The second question is whether different dimers affect the *positional* short range order at a length scale greater than one molecular length scale. This was studied through two methods: (a) the partial radial distribution functions, (b) the 2D positional maps like the one in figure 5.1. Concerning (a), three partial radial distributions were compared: the g_{OO} of all water molecules, the g_{OO} of water dimers with a *cis* acceptor, and the g_{OO} of water dimers with a *trans* acceptor. The result is shown in figure 5.5, there is no appreciable difference between these cases. This means that the two dimers cannot create differences in the microscopic molecular density, and it is therefore unlikely that they are responsible for high/low density liquid water ordering. Concerning (b), three 2D positional maps $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}})$ for molecules beyond the first hydration shell were compared using the same groups as before. Figure 5.6 shows the positional maps for the first 12 neighbours and the oxygen-oxygen radial distribution $g_{OO}(r)$ and the running coordination number for oxygens $n_{OO}(r)$ is presented in figure Figure 5.7. The MCN is the same as $n_{OO}(r)$, given that each water molecule has only one oxygen atom. Again there was no appreciable difference, with the agreement being slightly less good at distances associated to the first minimum (around MCN=7) in the $g_{OO}(r)$ radial distribution function where the probability to find a water molecules is low.

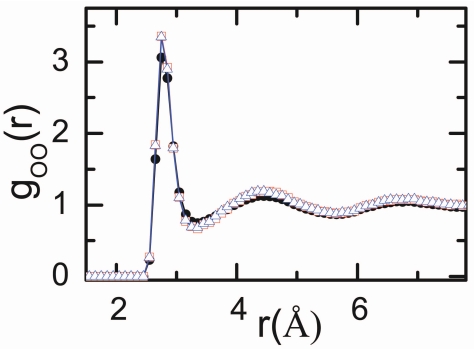

Figure 5.5: Total oxygen-oxygen radial distribution (black circles) and the partial radial distribution function between two oxygen atoms around molecules having a first neighbor either in cis (squares) or trans (triangles) configuration.

The third question is finally whether different dimers affect the *orientational* short range order at a length scale greater than one molecular length

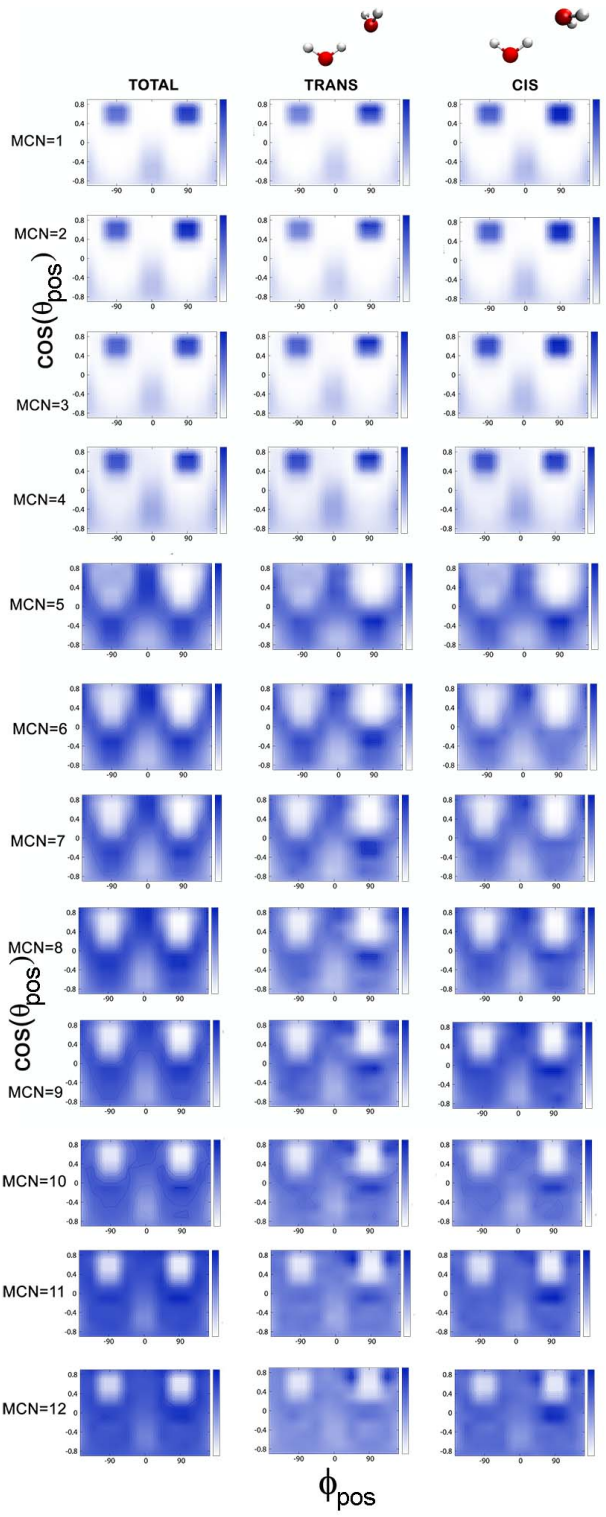

Figure 5.6: Positional maps $g(\cos(\theta_{\text{pos}}), \phi_{\text{pos}})$ for three cases: without dimer discrimination and being the neighbor molecule at $(\cos(\theta) \approx 0.66, \phi \approx 90°)$ *cis* or *trans*.

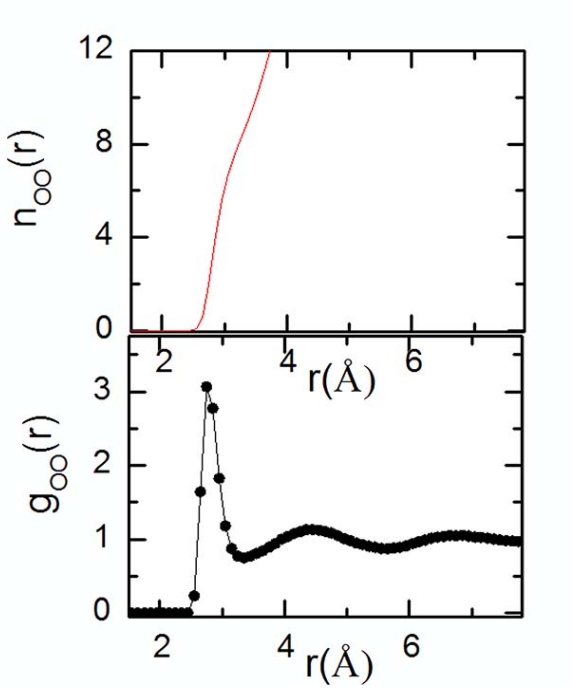

Figure 5.7: The partial radial distribution function $g_{OO}(r)$ and the molecular coordination number $n_{OO}(r)$.

scale. It is rather challenging to compare the orientational 3D probability distributions of Euler angles which depend on the molecular position and on the distance from the central molecule. For this reason it has been chosen to study the dipole moment orientation with respect to that of the central molecule ($\Phi_{\text{dip-dip}}$) as a characteristic molecular direction. This choice is also supported by the fact that, as it was demonstrated, differences in energies are caused by different dipole orientations. Although different dimers seem to have no effect on the position of water molecules and on the orientation in the first hydration shell, a different scenario for the orientation of molecules beyond the length scale of one molecule was found. Figure 5.8 shows the probability distribution of $\Phi_{\text{dip-dip}}$ as a function of the distance, starting from the first minimum of the g_{OO} partial radial distribution function, i. e. beyond the first hydration shell. Again, these calculations were performed for the three groups: all water molecules, the water dimers with a *cis* acceptor, and the water dimers with a *trans* acceptor. Figure 5.8 re-

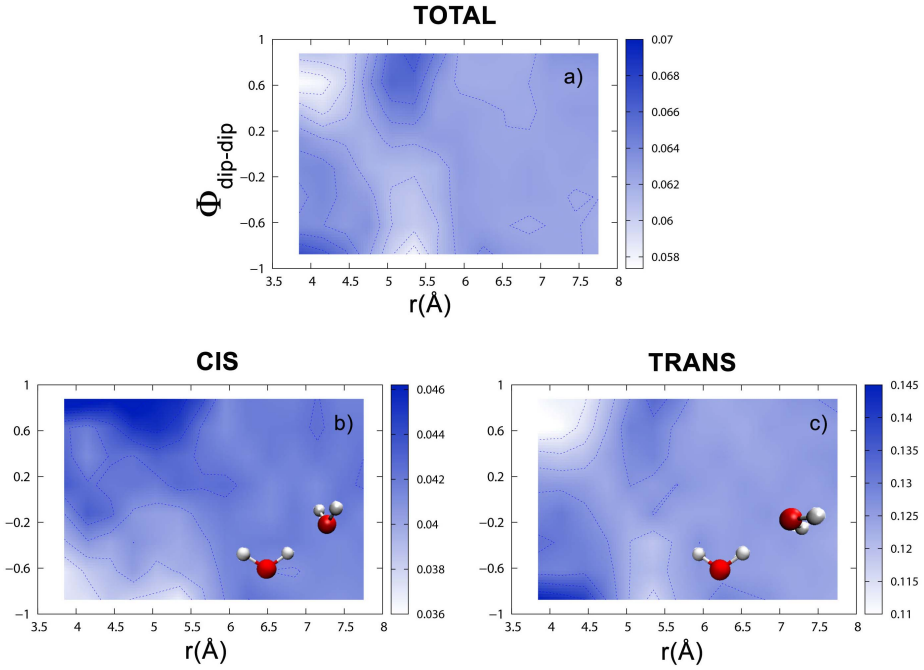

Figure 5.8: Probability map of the relative dipole orientation of a neighbour molecule with respect to the central one ($\Phi_{\text{dip-dip}}$) as a function of the distance, for molecules beyond the first hydration shell. Panel a shows the results for any water molecule,panels b and c show the calculations for molecules having an acceptor in *cis* or *trans* configuration respectively.

veals that the dipole orientation distribution for molecules is different up to distances of about 6 Å in the three cases. In particular, molecules in *cis* configuration (a parallel dipole moment) influence next neighbours so that they also align parallel up to a distance of about 6 Å, forming some kind of chain of dipoles. On the other hand molecules in a *trans* configuration alternate the dipole orientation also up to a distance of 6 Å.

5.6 Summary and conclusions

It is hitherto thought that liquid water is composed of tetrahedrally coordinated molecules with an asymmetric interaction of the central molecule with neighboring molecules. Kühne *et al., Nat. Commun.*, 2013, **4**, 1450 suggested that this asymmetry, energetic rather than geometric, is the cor-

nerstone to reconcile the homogeneous and inhomogeneous viewpoints of liquid water. In order to investigate the geometric origin of that asymmetry, Molecular Dynamics (MD) simulations of water have been scrutinized through a careful analysis of the five-dimensional probability distribution function of Euler angles in which the relative positions and orientations of water molecules are obtained. The results of this chapter demonstrate that, beyond the ubiquitous tetrahedral structure with well-defined molecular dimers, there is a series of possible molecular orientations that define the structure. These orientations are generated by rotating the neighboring molecule around the O-H axis that is involved in the hydrogen bond scheme. Two of the possible orientations have a higher probability, giving rise to two kinds of dimers: one close to the lowest energy of a water dimer in vacuum with an almost perpendicular alignment of the dipole moment, and another one with a parallel orientation of the dipole moment which is less tightly bound. These two different dimers have an effect on the orientation of further water dipole moments up to a distance of ≈6Å. Liquid water can therefore be described as a continuous mixture of two kinds of dimers where the hydrogen bonds have the same geometry but the interaction energies are different due to a different mutual orientation of the dipoles of the participating water molecules.

The structure of liquid water is therefore the result of a continuous mixture of molecules participating in one of two dimers with different energies that result from a different dipole orientation: on the one hand *trans* dimers that have almost perpendicular dipole moments and a low energy and on the other hand *cis* dimers that have parallel dipole moments and are less tightly bonded.

5.7 Bibliography

[1] J. L. Abascal and C. Vega. A general purpose model for the condensed phases of water: Tip4p/2005. *The Journal of chemical physics*, 123 (23):234505, 2005. 84

[2] H. Berendsen, J. Grigera, and T. Straatsma. The missing term in effective pair potentials. *Journal of Physical Chemistry*, 91(24):6269–6271, 1987. 84

[3] V. Buch, P. Sandler, and J. Sadlej. Simulations of h2o solid, liquid, and clusters, with an emphasis on ferroelectric ordering transition in hexagonal ice. *The Journal of Physical Chemistry B*, 102(44):8641–8653, 1998. 89, 90

[4] A. De Santis and D. Rocca. The local order in liquid water studied through restricted averages of the angular correlation function. *The Journal of chemical physics*, 107(22):9559–9568, 1997. 85

[5] D. S. Eisenberg, W. Kauzmann, et al. *The structure and properties of water*, volume 123. Clarendon Press Oxford, 1969. 89

[6] C. H. *et al. PNAS*, 106:15214–15218, 2009. 83

[7] P. W. *et al. Science*, 304:995–999, 2004. 83

[8] T. T. *et al. Chem. Phys. Lett.*, 460:387–400, 2008. 83

[9] M. A. González and J. L. Abascal. A flexible model for water based on tip4p/2005. *The Journal of chemical physics*, 135(22):224516, 2011. 84

[10] T. Head-Gordon and G. Hura. *Chem. Rev.*, 102:2651, 2002. 83

[11] T. D. Kühne and R. Z. Khaliullin. *Nature Communications*, 4:1450, 2013. 83

[12] T. D. Kuhne and R. Z. Khaliullin. Nature of the asymmetry in the hydrogen-bond networks of hexagonal ice and liquid water. *Journal of the American Chemical Society*, 136(9):3395–3399, 2014. 83, 85, 91

[13] A. Nilsson and L. G. M. Pettersson. *Chem. Phys.*, 389:1–34, 2011. 83

[14] L. Pusztai, O. Pizio, and S. Sokolowski. Comparison of interaction potentials of liquid water with respect to their consistency with neutron diffraction data of pure heavy water. *The Journal of chemical physics*, 129(18):184103, 2008. 84

[15] A. Soper. Orientational correlation function for molecular liquids: the case of liquid water. *The Journal of chemical physics*, 101(8):6888–6901, 1994. 85

[16] A. K. Soper. *Pure Appl. Chem.*, 82:1855, 2010. 83

[17] F. H. Stillinger. *Science*, 209:451–457, 1980. 83

[18] I. Svishchev and P. Kusalik. Structure in liquid water: A study of spatial distribution functions. *The Journal of chemical physics*, 99(4): 3049–3058, 1993. 85

[19] C. Vega and J. L. Abascal. Simulating water with rigid non-polarizable models: a general perspective. *Physical Chemistry Chemical Physics*, 13(44):19663–19688, 2011. 84

Chapter 6

The structure of liquid water beyond the first hydration shells

6.1 Motivation

Despite the huge efforts concerning the short-range order of liquid water, little attention has been paid on how this particular arrangement affects the structure of water molecules at larger distances. The work of Liu et al.(12) is an exception. They found angular correlations persisting on a length scale of ∼40 Å using Density Functional Theory. In that work it seemed that the special molecular arrangements are not correlated to the local density probed by the Oxygen-Oxygen partial radial distribution function $g_{\mathrm{OO}}(r)$. Higo et al. also found angular correlations in bulk water (4). They obtained a dipole chain order for distances up to ∼10 Å.

In this chapter, the distance-dependent excess entropy is used in order to characterize the structure of liquid water at distances of up to ∼10 Å. The excess entropy of a liquid is a way to quantify how different its structure is from the one of an ideal gas, thus it is defined as $S^{\mathrm{excess}} = S^{\mathrm{liq}} - S^{\mathrm{gas}}$ where S^{liq} and S^{gas} are the entropies related to the liquid and the ideal gas. Since the ideal gas is more disordered than the liquid, this magnitude is always

negative. The calculation of the excess entropy involves the determination of the N-body correlation function that specifies the position of all the particles in the liquid.

6.2 Computational details

6.2.1 Simulation details

Molecular Dynamics (MD) simulations were performed in order to characterize the long range order in liquid water. The simulations were carried out using the Gromacs (Groningen Machine for Chemical Simulation) engine (7) along with the TIP4P/2005 (1) forcefield as in the previous chapter. This model for water was chosen since it has been found that it is the best available to describe the structural properties of liquid water when compared with many other water models (18, 20, 23). Moreover the TIP4P/2005 water model is also the best model to reproduce many other experimental quantities (1) at ambient conditions. The time step was chosen as 1 fs since it is a rigid model, constraining the angles and bonds with the LINCS algorithm (3). The simulation was run on 432 molecules in an NPT ensemble at the thermodynamic conditions of liquid water, T=298 K and P=1 bar using the Nose-Hoover (6) thermostat and an isotropic Parrinello-Rahman (17) barostat with coupling relaxation times of 2 ps. The cut-off used was 9.0 Å for the van der Waals and 8.5 Å for the short-range electrostatic interactions; the particle mesh Ewald method was used beyond the electrostatic cut-off for the reciprocal space sum. 5000 snapshots of the simulation were saved for further analysis every one picosecond.

6.2.2 Axes definition

Figure 6.1 shows the definition of the angles used to analyze the MD trajectories: θ_{pos} and ϕ_{pos} describe the position of the second molecule in spherical coordinates chosing the dipolar axis of water as z axis and the x axis perpendicular to the plane of the molecule. The Euler angles θ_{ori} and ϕ_{ori} describe the orientation of the dipolar axis of the second molecule with respect to the central one, and ψ_{ori} describes the rotation of the second water molecule around that axis.

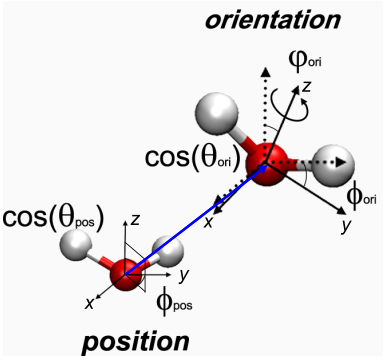

Figure 6.1: Definition of axes used to calculate the positional $g\left(\theta_{\text{pos}}, \phi_{\text{pos}}|r\right)$ and orientational $g\left(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}|r\right)$ correlation functions.

6.3 Structure characterization by excess entropy calculation

Information theory is used in this chapter in order to be able to evaluate the wealth of information contained in the trajectories of the molecules in an MD simulation. It provides a framework to extract useful information about the system by looking at the different contributions to its excess entropy. The excess entropy of any liquid S_{excess}, i. e. the difference to the entropy of an ideal gas at the same average density and temperature, can be calculated from the six-dimensional correlation function $g(r, \Omega)$ as a function of the distance r and a group of five angles defining the relative position and orientation Ω. When restricted to a given length scale (a determined distance r between particles) $g(r, \Omega)$ depends only on five angular variables, and the correlation function will be then written as $g(\Omega|r)$. A good approximation for the calculation of the excess entropy is to consider only the term related to two-particle correlations. Other authors (11, 23) have shown that two-body correlation are able to account for 85%-95% of excess entropy in the case of simple hard sphere Lennard-Jones fluids. In this investigation, two-body correlations functions were used since previous works have shown that even if water is certainly not a hard-sphere liquid, the calculation using two-body correlations is able to give a value for excess entropy very close to the experimental one (11).

The excess entropy can be then separated into two terms (10)

$$S_{\text{excess}} = S^{\text{tot}}_{\text{trans}} + S^{\text{tot}}_{\text{ang}} \tag{6.1}$$

where $S^{\text{tot}}_{\text{trans}}$ is the total entropy connected to the particle density at a given length scale, i. e. features in $g_{OO}(r)$, and $S^{\text{tot}}_{\text{ang}}$ contains all contributions that are a function of Ω. These two contributions can be analyzed as functions of the distance r, defined as

$$S^{\text{tot}}_{\text{trans}} = \frac{1}{2} k\rho \int S_{\text{trans}}(r)\ \mathrm{d}r \tag{6.2}$$

$$S_{\text{trans}}(r) = -\left[g_{OO}(r) \ln g_{OO}(r) - g_{OO}(r) + 1\right] 4\pi r^2 \tag{6.3}$$

with the Boltzmann factor k and number density ρ, and

$$S^{\text{tot}}_{\text{ang}} = \frac{1}{2} k\rho \int 4\pi r^2 g_{OO}(r) S_{\text{ang}}(r)\ \mathrm{d}r \tag{6.4}$$

$$S_{\text{ang}}(r) = -\frac{1}{\Omega} \int g(\Omega|r) \ln g(\Omega|r)\ \mathrm{d}\Omega \tag{6.5}$$

$$= S_{\text{pos}}(r) + S_{\text{ori}}(r) + S_{\text{pos}\star\text{ori}}(r) \tag{6.6}$$

where $g(\Omega|r)$ denotes the five-dimensional Ω distribution function of molecules at a given value of the distance r. As shown in the equation, it is possible to write this quantity as a sum of three parts, S_{pos} stemming from the relative position of two molecules $(\theta_{\text{pos}}, \phi_{\text{pos}})$, S_{ori} stemming from their relative orientation $(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}})$, and the cross-term $S_{\text{pos}\star\text{ori}}$.

The expansion of the correlation function can be performed, so that the entropy $S_{\text{ang}}(r)$ can be calculated as the entropy $S^{(1)}(r)$ of one-angle correlation functions $g(\alpha|r)$, corrected by the contribution of correlation functions of increasing order; $S^{(2)}(r)$ from $g(\alpha, \beta|r)$; $S^{(3)}(r)$ from $g(\alpha, \beta, \gamma|r)$ and so on. With this method, the entropy $S^{(5)}(r)$ of the five-dimensional correlation function appearing in equation 6.5 can be calculated as the summation of the entropies related to each angle, corrected by higher-order mutual information terms I (8, 13, 14) which here are written out up to the third

order and neglect fourth-order terms:

$$S^{(5)}(r) = \sum S^{(1)}(r) - \sum I^{(2)}(r) + \sum I^{(3)}(r) + \mathcal{O}^{(4)} \tag{6.7}$$

where

$$I^{(2)}(r) = \sum S^{(1)}(r) - S^{(2)}(r) \tag{6.8}$$

$$I^{(3)}(r) = \sum S^{(1)}(r) - \sum S^{(2)}(r) + S^{(3)}(r) \tag{6.9}$$

so that for example the mutual information of two variables A and B, $I(A, B)$ of type $I^{(2)}$, is calculated as $I(A, B) = S(A) + S(B) - S(A, B)$.

This expansion is firmly based on information theory concepts (13) and it is a most powerful method to study the ordering of molecules in disordered phases (14). Mutual information is a measure of the interdependency of N variables: high mutual information values in a two-fold correlation function imply that the two variables are strongly correlated, and therefore that a factorization is not possible ($g(\alpha, \beta) \neq g(\alpha)g(\beta)$). On the contrary a low value of mutual information implies that variables are independent, allowing a factorization of the correlation function. If the excess entropy is expanded up to the third order it is possible to extract valuable physical information from the expansion of equation (6.7).
The variables determining the position of two neighboring water molecules, at a certain distance, in spherical coordinates (θ_{pos} and ϕ_{pos}) can be separated from the Euler angles related to their relative orientation (θ_{ori}, ϕ_{ori} and ψ_{ori}) (for a definition of the angles see figure 6.1). Separating the contributions due to the relative position and relative orientation of two molecules one obtains, following equation (6.8), for the position:

$$S_{\text{pos}}(\text{r}) := S(\theta_{\text{pos}}|r) + S(\phi_{\text{pos}}|r) - I(\theta_{\text{pos}}, \phi_{\text{pos}}|r) \tag{6.10}$$

or

$$S_{\text{pos}}(\theta_{\text{pos}}, \phi_{\text{pos}}|r) := -\frac{1}{\Omega_{\text{pos}}} \int g\left(\theta_{\text{pos}}, \phi_{\text{pos}}|r\right) \ln\left(g\left(\theta_{\text{pos}}, \phi_{\text{pos}}|r\right)\right) \mathrm{d}\Omega \tag{6.11}$$

and for the relative orientation, using equation (6.9)

$$S_{\text{ori}}(\text{r}) := S(\theta_{\text{ori}}|r) + S(\phi_{\text{ori}}|r) + S(\psi_{\text{ori}}|r) - I(\theta_{\text{ori}}, \phi_{\text{ori}}|r) - I(\theta_{\text{ori}}, \psi_{\text{ori}}|r) - I\left(\phi_{\text{ori}}, \psi_{\text{ori}}|\, r\right) + I(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}|r) \quad (6.12)$$

or

$$S_{\text{ori}}(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}|r) := -\frac{1}{\Omega_{\text{ori}}} \int g\left(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}|r\right) \ln g\left[(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}|r)\right] \mathrm{d}\Omega \quad (6.13)$$

and finally the cross-term between position and orientation in equation (6.6) can be written as:

$$\begin{aligned} S_{\text{pos}\star\text{ori}}(\text{r}) := S_{\text{ang}}(\text{r}) - S_{\text{pos}}(\text{r}) - S_{\text{ori}}(\text{r}) \\ = - \sum_{\substack{\omega_{\text{pos}} \in \theta_{\text{pos}}, \phi_{\text{pos}} \\ \omega_{\text{ori}} \in \theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}}} I(\omega_{\text{pos}}, \omega_{\text{ori}}) \\ + \sum_{\substack{\omega_{\text{pos}} \in \theta_{\text{pos}}, \phi_{\text{pos}} \\ \omega'_{\text{pos}} \in \theta_{\text{pos}}, \phi_{\text{pos}} \\ \omega_{\text{ori}} \in \theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}}} I(\omega_{\text{pos}}, \omega'_{\text{pos}}, \omega_{\text{ori}}) \\ + \sum_{\substack{\omega_{\text{pos}} \in \theta_{\text{pos}}, \phi_{\text{pos}} \\ \omega_{\text{ori}} \in \theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}} \\ \omega'_{\text{ori}} \in \theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}}} I(\omega_{\text{pos}}, \omega_{\text{ori}}, \omega'_{\text{ori}}) \\ + \mathcal{O}^{(4)} \end{aligned} \quad (6.14)$$

Where $\Omega_{\text{pos}} = 4\pi$ is the integral over the angles defining the position θ_{pos} and ϕ_{pos}, and $\Omega_{\text{ori}} = 8\pi^2$ is the integral over the angles defining the orientation θ_{ori}, ϕ_{ori} and ψ_{ori}. In these equations, the angular position of a second molecule at a certain distance r is encoded in the correlation function $g\left(\theta_{\text{pos}}, \phi_{\text{pos}}|r\right)$, and their relative orientation in $g\left(\theta_{\text{ori}}, \phi_{\text{ori}}, \psi_{\text{ori}}|r\right)$. The aforementioned separation is justified by analyzing how strongly correlated the angular variables are, via an analysis of the mutual information between them shown in figure 6.2. This separation into purely positional and orien-

tational contributions is only possible when the entropy is expanded up to the third order, for higher orders the variables are inevitably mixed.

It must be highlighted that correlation functions involving an angle between vectors α such as θ_{pos} and θ_{ori} have been calculated using a non-uniform binning in angle (cosinus of the angle). This is necessary in order to have a featureless correlation function when two vectors are randomly distributed. If the angle is used to perform the binning, the correlation function between two vectors would have a maximum at $\alpha = 90°$ and would be zero at $\alpha = 0°$ and $\alpha = 180°$. This would give rise to non-vanishing mutual information terms that would alter the values of the total excess entropy.

According to the axes set shown in figure 6.1, θ_{ori} is therefore related to the relative orientation of two water dipoles, and ϕ_{ori} and ψ_{ori} to the rotation around this axis. The analysis was carried out with the program ANGULA (16), choosing an angular increment of 20° for angles between planes and an increment for the cosine of angles between vectors of 0.2. Excess entropies, and therefore also mutual information terms, have been rescaled so that they vanish for long distances (5). The distance increment to produce each map has been chosen to be Δr=0.1 Å. The excess angular entropy obtained with these conditions is 10.34 cal/K/mol (86% of the total excess angular entropy) in good agreement with literature values (8). Spatial Distribution Maps were plotted using VMD (9).

Figure 6.2 shows the integrated second order mutual information contributions to the total entropy, in other words: correlations between different pairs of angles. They have been calculated similarly to the total entropy in equation (6.2), i. e. $I^{(i)\mathrm{tot}} = -\frac{1}{2}k\rho \int I^{(i)}(r)\mathrm{d}r$. The strongest correlation between two angles is not surprisingly the one relating the positional angles θ_{pos} and ϕ_{pos}. Concerning the purely orientational contributions, the correlation between the variables ϕ_{ori} and ψ_{ori} is stronger than the correlations between any of these angles and the one defining the dipole orientation θ_{ori}. The correlation function of the relative orientation of two molecules

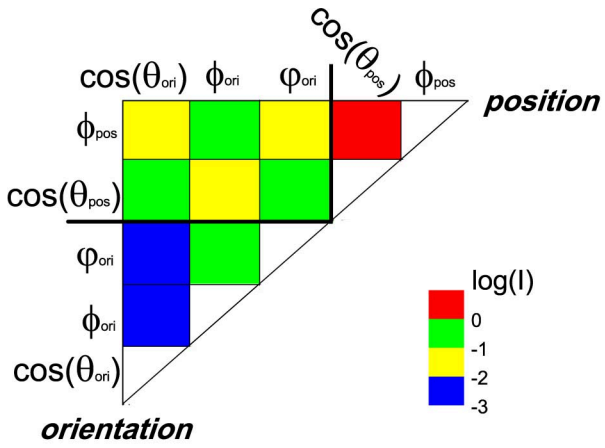

Figure 6.2: Logarithm of the integrated mutual information values calculated for all possible combinations of angles or their cosine (cf. text). The angles have been grouped separating those describing the relative position of two molecules (upper right) from the orientation (lower left).

can therefore be naturally factorized into one term concerning the dipole orientation and another term defining the rotation around the dipolar axis: $g(\theta_{\mathrm{ori}}, \phi_{\mathrm{ori}}, \psi_{\mathrm{ori}}) \sim g(\theta_{\mathrm{ori}}) \cdot g(\phi_{\mathrm{ori}}, \psi_{\mathrm{ori}})$.

6.4 Results for the excess entropy contributions

Figure 6.3 shows the different contributions to the excess entropy along with the radial distribution function $g_{\mathrm{OO}}(r)$. As it can be readily seen figure 6.3, the maxima of the translational excess entropy contribution line up with extrema (maxima and minima) of $g_{\mathrm{OO}}(r)$. This is not surprising as this contribution is directly connected to the local density, cf. equation (6.3). On the other hand the contribution coming from the crossed term $S_{\mathrm{pos}\star\mathrm{ori}}$ (cf. equation 6.14) does not provide any information because it is almost featureless.

The interesting contributions depicted in figure 6.3 are $S_{\mathrm{pos}}(r)$ and $S_{\mathrm{ori}}(r)$ whose calculation *is not connected to deviations of the local den-*

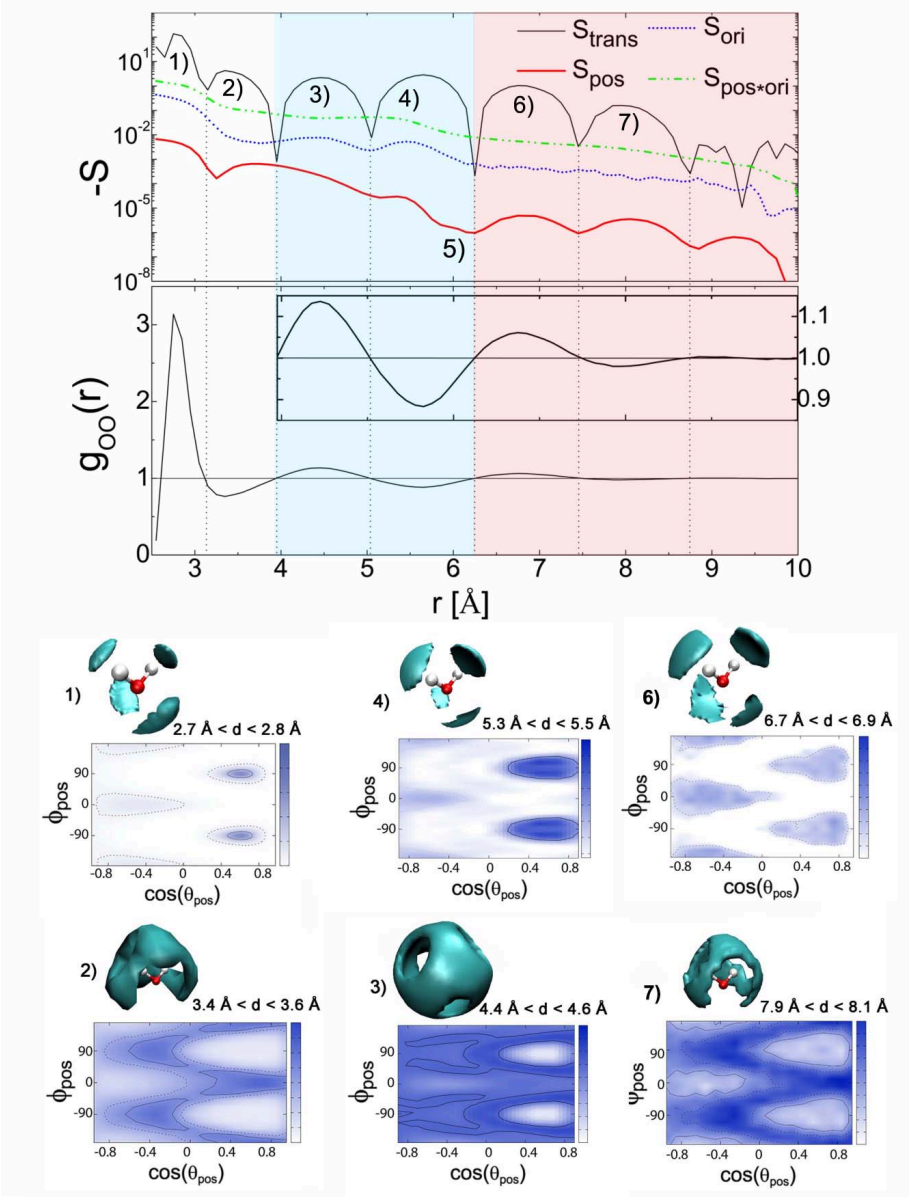

Figure 6.3: Upper panel: Different contributions to the entropy S as a function of distance on a logarithmic scale. Black thin line: Absolute translational contribution $S_{\mathrm{trans}}(r)$; red thick line: positional contribution $S_{\mathrm{pos}}(r)$; blue dotted line: orientational contribution $S_{\mathrm{ori}}(r)$; green dash-dotted line: positional and orientational contributions $S_{\mathrm{pos}\star\mathrm{ori}}(r)$ to the total third-order entropy. Also shown in the lower panel is the oxygen-oxygen radial distribution function, with a zoom in the inset to show the oscillations more clearly. Lower panel: Spatial density maps and distribution functions $g\left(\theta_{\mathrm{pos}}, \phi_{\mathrm{pos}} | r\right)$ associated to the regions defined in figure 6.3: (1) the first four neighbours: 2.7-2.8 Å, (2) second four nearest neighbours: 3.4-3.6 Å, thereafter, molecules at distances (3) 4.4-4.6 Å (4) 5.3-5.5 Å (5) 6.7-6.9 Å and (6) 7.9-8.1 Å.

sity probed by $g_{OO}(r)$. After the first two hydration shells, i. e. at distances between 4.0 Å and 6.2 Å, there is a correlation between the value of the translational entropy (and thus the $g_{OO}(r)$) and the purely *orientational* contribution to the excess entropy $S_{ori}(r)$: the maxima of both functions are located at the same distances. For longer distances the orientational contribution is structureless and simply decays as the system gets more disordered. On the contrary, for these distances (larger than 6.2 Å), it is the *positional* contribution to the excess entropy $S_{pos}(r)$ which is correlated with the translational entropy and therefore also with the $g_{OO}(r)$: both functions have the same series of features for these distances. It is worth noting that $S_{pos}(r)$ displays a small peak at ca. 5.4 Å, meaning that at these length scales there is a smooth transition from the orientational regime to the positional one.

Starting with the features in $S_{pos}(r)$, the aim is to find out which molecular arrangements are giving rise to the maxima of the positional entropy at $r > 6.2\,\mathring{A}$. For this purpose, in the Lower panel of figure 6.3 the two-dimensional correlation function $g(\theta_{pos}, \phi_{pos}|r)$ was generated at distances ranged within $6.7\,\mathring{A} < r < 6.9\,\mathring{A}$ (region 6) and $7.9\,\mathring{A} < r < 8.1\,\mathring{A}$ (region 7) where the maxima are observed. Also plotted in the same figure are the results for the first (region 1) and second (region 2) regions, i. e. the first four water molecules sitting in a tetrahedron and the next four sitting in the vacancies of the first tetrahedron. For the sake of completeness, in the same figure the results of the third and the fourth hydration shells are also shown. Also shown in figure 6.4 the results for all distances in steps of 0.1 Å. It can be seen in that figure, in agreement with previous works, that molecules continue to fill the gaps of previous hydration shells in a more or less disordered way (15, 19, 22). From all the maps calculated in figure 6.4, there are shown those for region 3, where there is no clear maximum in the $S_{pos}(r)$ function, and those for region 4 where a first bump is seen in $S_{pos}(r)$. In agreement with the shape of $S_{pos}(r)$ function, for region 3 (where $S_{pos}(r)$ is flat) there is not a clear structure, and for region 4 (where a first maxima in $S_{pos}(r)$ is to be seen) the molecular ordering is more clear. The reader must be warned that the regions have been chosen to correspond with maxima in

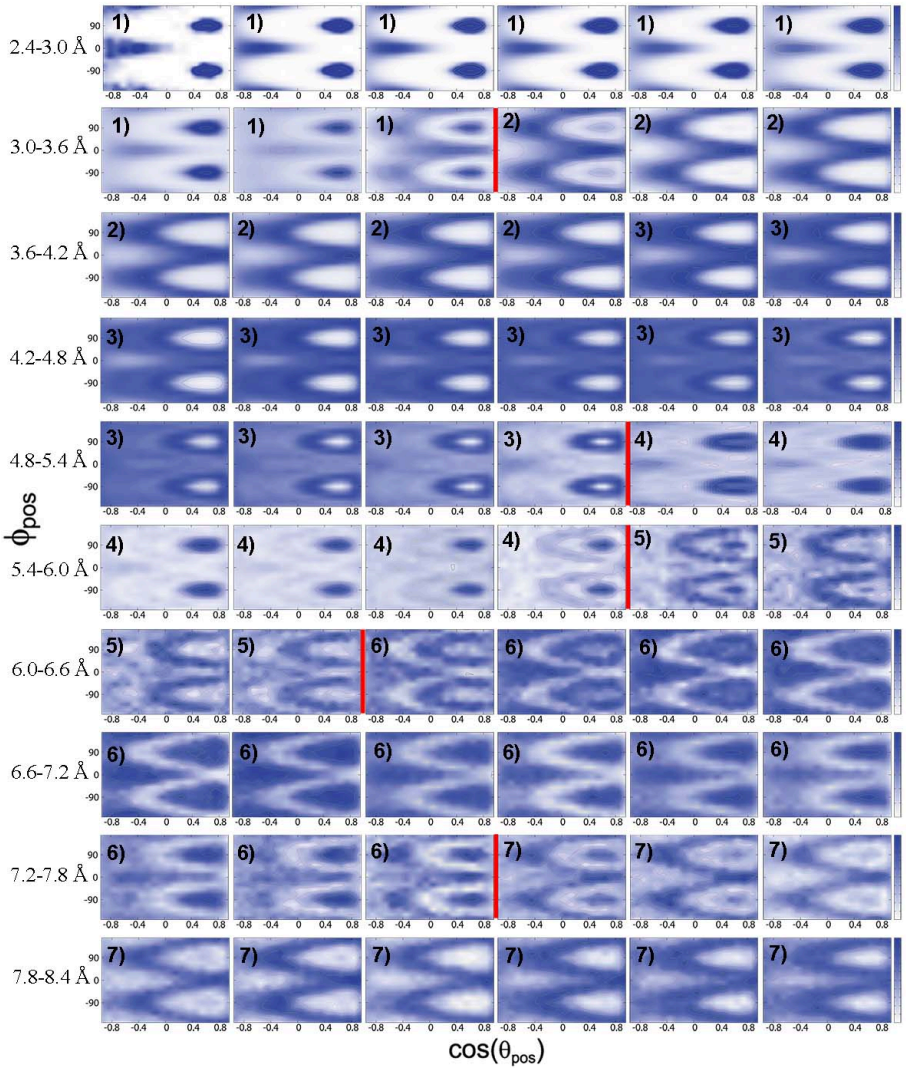

Figure 6.4: (color online). Distribution functions $g\left(\theta_{\mathrm{pos}}, \phi_{\mathrm{pos}}|r\right)$ for the oxygen atom in liquid water. 60 maps have been generated, each one within a 0.1 Å from the minimum distance 2.4 Å to 8.4 Å. The first map corresponds with a distance range of 2.5-2.5 Å, the second map to 2.5-2.6 Å, and so on until the 60th map within a range of 8.3-8.4 Å. Blue colors represent maxima in probability while white colours represent probability minima.

$S_{\text{trans}}(r)$ and not to successive shells where the alternance of shells is to be seen. This can be seen in figure 6.4. To make the meaning of $g(\theta_{\text{pos}}, \phi_{\text{pos}})$ more intuitive, the same information is shown as three-dimensional Spatial Distribution Maps (21) (SDM). It is immediately visible both looking at the correlation function and the generated SDM for region 6 that the structure is astonishingly well defined at these long distances. Even more surprising is that molecules at such long distances reproduce the tetrahedral structure of the first two hydration shells at short distances. The maps show unambiguously that the structures are much better defined than one might have expected from the small values of the excess entropy. Therefore, there is a correlation between an angular magnitude and both the $g_{\text{OO}}(r)$ and $S_{\text{trans}}(r)$, contrary to the results of earlier studies (12).

6.5 The orientational contribution $S_{\text{ori}}(r)$

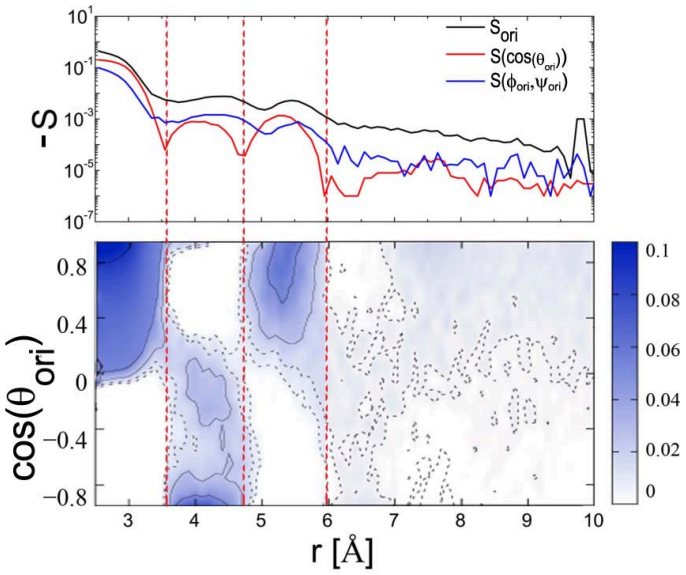

Figure 6.5: Upper panel: Orientational entropy S_{ori} and the two independent contributions $g(\cos(\theta_{\text{ori}}))$ and $g(\phi_{\text{ori}}, \psi_{\text{ori}})$ as a function of distance. Lower panel: Probability distribution function of the cosine of the angle between dipoles $\cos(\theta_{\text{ori}})$ as a function of the distance. Dashed lines serve as an eye guide to delimitate regions of maxima and minima arising from $g(\cos(\theta_{\text{ori}}))$.

The features in $S_{\text{ori}}(r)$ correlate in the distance range between 4.0 Å and

6.2 Å with the maxima and minima of the translational entropy. In order to investigate the origin of this correlation, the orientational contribution was further separated into two independent contributions, $g(\cos(\theta_{\mathrm{ori}}))$ and $g(\phi_{\mathrm{ori}}, \psi_{\mathrm{ori}})$, which correspond to the angle between the two dipoles and the rotation around this axis, respectively. Their contributions to the orientational excess entropy can be calculated as:

$$S\left(\cos(\theta_{\mathrm{ori}})\right) = -\frac{1}{2}\int g\left(\cos(\theta_{\mathrm{ori}})\right) \ln\left[g\left(\cos(\theta_{\mathrm{ori}})\right)\right] \mathrm{d}\cos(\theta_{\mathrm{ori}}) \tag{6.15}$$

$$S\left(\phi_{\mathrm{ori}}, \psi_{\mathrm{ori}}\right) = -\frac{1}{4\pi^2}\int g\left(\phi_{\mathrm{ori}}, \psi_{\mathrm{ori}}\right) \ln\left[g\left(\phi_{\mathrm{ori}}, \psi_{\mathrm{ori}}\right)\right] \mathrm{d}\phi_{\mathrm{ori}}\mathrm{d}\psi_{\mathrm{ori}} \tag{6.16}$$

Figure 6.5 shows the values of the orientational excess entropy $S_{\mathrm{ori}}(r)$ together with the two parts arising from these two correlation functions. They rule the behaviour of the total orientational excess entropy S_{ori}, exhibiting all features at about the same distances for 4.0 Å¡r¡6.2 Å.

Since the angle between dipoles, and the rotation around this axis, seem to be the main contributions to orientational excess entropy, figure 6.5 shows the probability distribution function of the cosine of the dipolar angle as a function of the distance. It shows the existence of an alternating dipole orientation starting from being parallel in the first hydration shell (16), turning to an antiparallel alignment and back to being parallel before vanishing in noise at r ∼6 Å. Such a spontaneously formed chain of dipoles might be related to the vortices that were found to form around molecules in aqueous solution which were suggested to be important for mediating interactions between the molecules over larger distances (2, 4).

6.6 Summary and conclusions

Excess entropy calculations allow to determine if changes in molecular ordering and in local density are correlated with the relative position or orientation of molecules. Using these functions it seems that the ordering of

liquid water at distances larger than ∼4 Å is ruled for distances up to about ∼6 Å by an alternating *orientation* of water dipoles, while at distances between ∼6 Å and ∼9 Å it is the *position* of water molecules that dominates the changes in molecular ordering. Even at distances as long as 9 Å there are "shells" which reproduce the well-known tetrahedrality of the first four water molecules.

This special positional order could constitute a new mechanism for long-distance interactions between molecules in aqueous solution, such as for example in molecular recognition or protein-protein interactions, which would play a role together with long dipole-chains reported in the literature(2, 4).

6.7 Bibliography

[1] J. L. F. Abascal and C. Vega. *J. Chem. Phys.*, 123:234505, 2005. 100

[2] A. N. Dickey and M. J. Stevens. *Phys. Rev. E*, 86:051601, 2012. 111, 112

[3] B. H. *et al. J. Comput. Chem.*, 18:1463–1472, 1997. 100

[4] J. H. *et al. PNAS*, 98:5961–5964, 2001. 99, 111, 112

[5] R. S. *et al. Bioinformatics*, 18:(suppl 2): S231–S240, 2002. 105

[6] D. J. Evans and B. L. Holian. *J. Chem. Phys.*, 83:4069–4074, 1985. 100

[7] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Comput.*, 4:435–447, 2008. 100

[8] D. J. Huggins. *J. Chem. Phys.*, 136:064518, 2012. 102, 105

[9] W. Humphrey, A. Dalke, and K. Schulten. *J. Molec. Graphics*, 14(1): 33–38, 1996. 105

[10] T. Lazaridis and M. Karplus. Orientational correlations and entropy in liquid water. *J. Chem. Phys.*, 105(10):4294–4316, 1996. 102

[11] T. Lazaridis and M. Karplus. *J. Chem. Phys.*, 105:4294, 1996. 101

[12] Y. Liu and L. Wu. *J. Chem. Phys.*, 139:041103, 2013. 99, 110

[13] H. Matsuda. *Phys. Rev. E.*, 62:3096–3102, 2000. 102, 103

[14] A. C. Pardo, A. Henao, and A. Vispa. *J. Non-Cryst.*, 407:220–227, 2015. 102, 103

[15] L. C. Pardo, N. Veglio, F. J. Bermejo, J. L. Tamarit, and G. J. Cuello. Experimental assessment of the extent of orientational short-range order in liquids. *Phys. Rev. B*, 72(1):014206, 2005. 108

[16] L. C. Pardo, A. Henao, S. Busch, E. Guàrdia, and J. L. Tamarit. *Phys. Chem. Chem. Phys.*, 16:24479–24483, 2014. 105, 111

[17] M. Parrinello and A. Rahman. *J. Appl. Phys.*, 52:7182, 1981. 100

[18] L. Pusztai, O. Pizio, and S. Sokolowski. Comparison of interaction potentials of liquid water with respect to their consistency with neutron diffraction data of pure heavy water. *The Journal of chemical physics*, 129(18):184103, 2008. 100

[19] M. Rovira-Esteva, A. Murugan, L. C. Pardo, S. Busch, M. D. Ruiz-Martin, M.-S. Appavou, J. L. Tamarit, C. Smuda, T. Unruh, and F. J. Bermejo. Microscopic structures and dynamics of high-and low-density liquid t r a n s-1, 2-dichloroethylene. *Phys. Rev. B*, 81(9):092202, 2010. 108

[20] Z. Steinczinger and L. Pusztai. Comparison of the tip4p-2005, swm4-dp and bk3 interaction potentials of liquid water with respect to their consistency with neutron and x-ray diffraction data of pure water. *Condensed Matter Physics*, 16(4), 2013. 100

[21] I. Svishchev and P. Kusalik. Structure in liquid water: A study of spatial distribution functions. *J. Chem. Phys.*, 99(4):3049–3058, 1993. 110

[22] N. Veglio, F. J. Bermejo, L. C. Pardo, J. L. Tamarit, and G. J. Cuello. Direct experimental assessment of the strength of orientational correlations in polar liquids. *Phys. Rev. E*, 72(3):031502, 2005. 108

[23] J. Zielkiewicz. Structural properties of water: Comparison of the spc, spce, tip4p, and tip5p models of water. *The Journal of chemical physics*, 123(10):104501, 2005. 100, 101

Chapter 7

Short range order and dynamics in the plastic crystal phases of water.

7.1 Motivation

In recent years it has been found by means of molecular dynamics simulations the appearance of plastic crystal phases of water. Aragones and Vega found in 2008 a bcc and fcc plastic crystal phases trying to study the nucleation to ice vii (1, 2). Figure 7.1 shows how the phase diagram would look for the TIP4P/2005 water model including those phases. In a parallel work, Takii and coworkers found the bcc plastic crystal phase (8) and they studied also the dynamics of that plastic crystal (5). These findings can be of great importance: on the one hand they can prove the validity of simulation and computer models of water and its power to predict experimental results if found a posteriori, on the other hand the non-existence of such plastic phases would mean that a redirection on the models of water should be taken. In either case, it is a step forward for computational research and the development of improved water models.

In this chapter, the bcc and fcc plastic phases of water are reproduced and the analysis tool developed in this thesis have been applied to under-

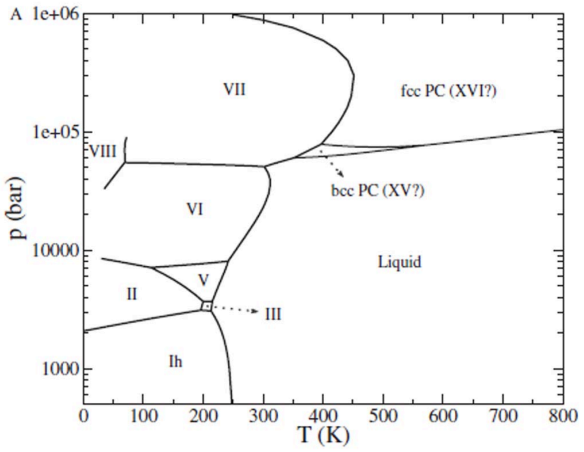

Figure 7.1: Global phase diagram for the TIP4P/2005 model including the bcc and fcc plastic crystals. *Modified from Ref. (2)*

stand better their structure and dynamics compared to the liquid and ice vii phases.

7.2 Computational details

7.2.1 From liquid water to the plastic crystal phases

The plastic crystal phases of water are found in the high pressure range, around 8 GPa. In this section the simulations of Aragones et al. (1, 2) have been reproduced to obtain the bcc and fcc plastic phases and also a high pressure liquid and the ice vii crystal. The simulation details are shown in detail in this section in order to introduce the characteristics of each phase, figure 7.2 shows a summary for the obtention of the different water phases, it is also explained carefully next.

The molecular dynamics parameters, such as the cut-off, thermostat and barostat were chosen as in chapter 6, see subsection 6.2.1 for details. The configurations were obtained by starting from a randomly disordered system of 432 water molecules at ambient conditions using a cubic box. The liquid water obtained was subsequently cooled by pressure. Serial simulations were ran at T = 440K increasing the pressure in steps of one gigapascals,

each simulation was equilibrated and the final configuration was taken as the initial configuration for the next step. The result was a high pressure liquid at T = 440K, P = 8GPa. This is a metastable state of water as it lies within the region of the bcc plastic crystal in the phase diagram (1)). The liquid was simulated under constant temperature and pressure for 5 ns to see the liquid-bcc transition reported (1, 2, 8). The density evolution is shown for the liquid-bcc transition in figure 7.2. There was a spontaneous change in pressure around 2 ns, from 1626 kg/m^3 to 1661 kg/m^3. The snapshot of the bcc plastic crystal shows that it has positional order but the molecules are orientationally disordered.

The new bcc plastic crystal was simulated afterwards at a constant pressure of 7 GPa and the temperature was lowered in steps of 5 K in the range from 440 K to 330 K allowing the system to equilibrate after each cooling step. The transition to the crystal ice vii was localized at $T = 365\,K$: $1625\,\mathrm{kg/m^3} \rightarrow 1688\,\mathrm{kg/m^3}$. The temperature annealing and the density evolution are shown for the bcc-ice vii transition in figure 7.2. The snapshot of the ice vii shows a positional arrangement similar to the bcc phase but an increased orientational order corresponding to a crystal phase.

For the remaining fcc plastic crystal an MD simulation was done in the (T = 460K, P = 8GPa) point of the phase diagram found by Aragones and Vega(1). The fcc lattice was obtained with 500 water molecules[125 unit cells (5x5x5)] equally orientated. Random orientations were assigned by running a short simulation with position restraints applying a large energy penalty for the displacements. Afterwards, a regular MD was ran for 5ns where the fcc PC remained stable, see figure 7.3. Although the fcc plastic crystal was not obtained from random orientations as it was the case of the bcc and ice vii phases, Aragones and Vega (1) showed that the final structure of the fcc plastic crystal was identical between two simulations using a random and fcc lattice for the initial configurations.

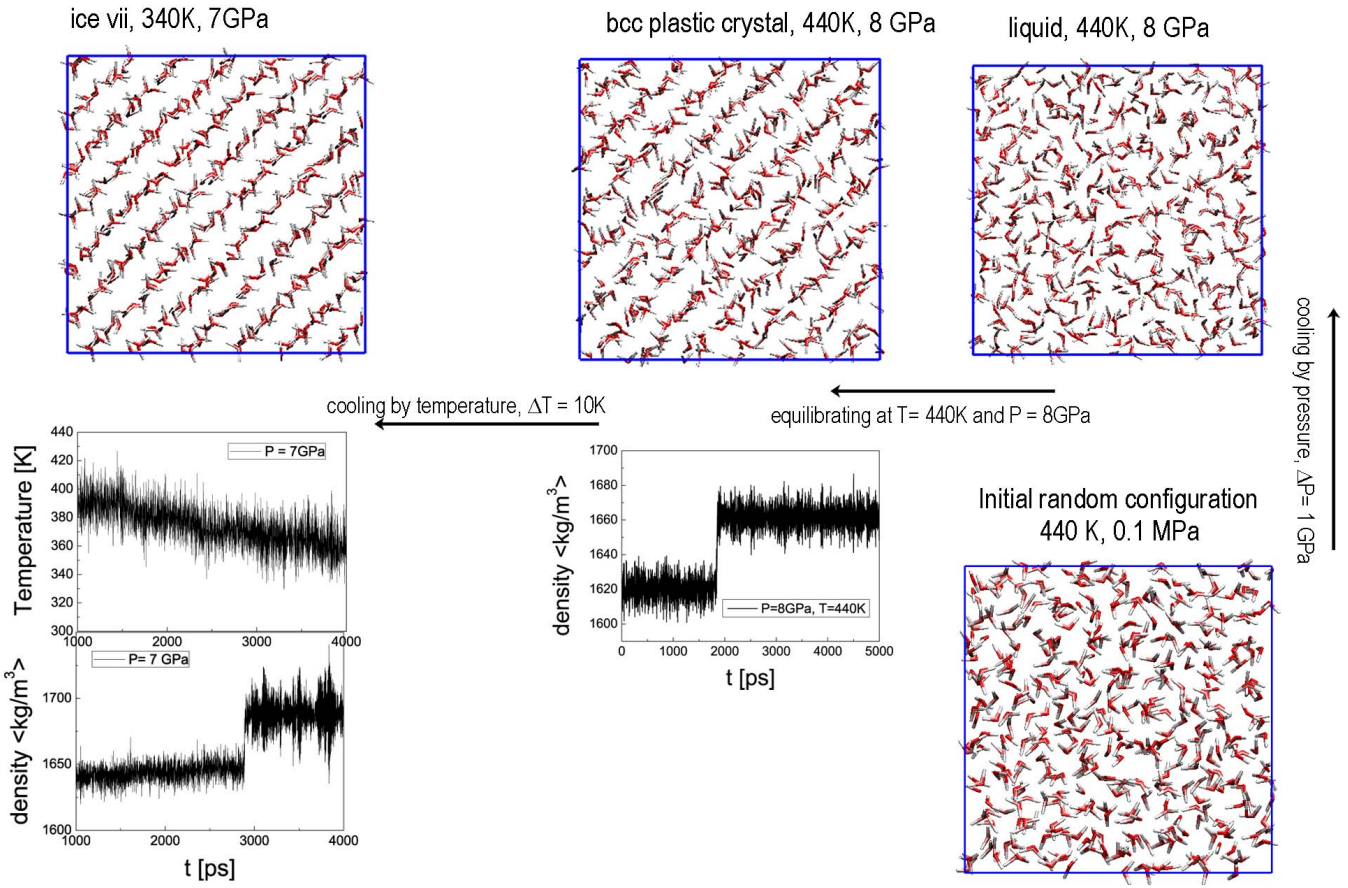

Figure 7.2: Diagram showing the phase transition between the liquid, bcc plastic crystal and ice vii phases obtained using molecular dynamics. The inset figure was taken from (2) and it shows the TIP4P/2005 phase diagram including the bcc and fcc plastic crystals.

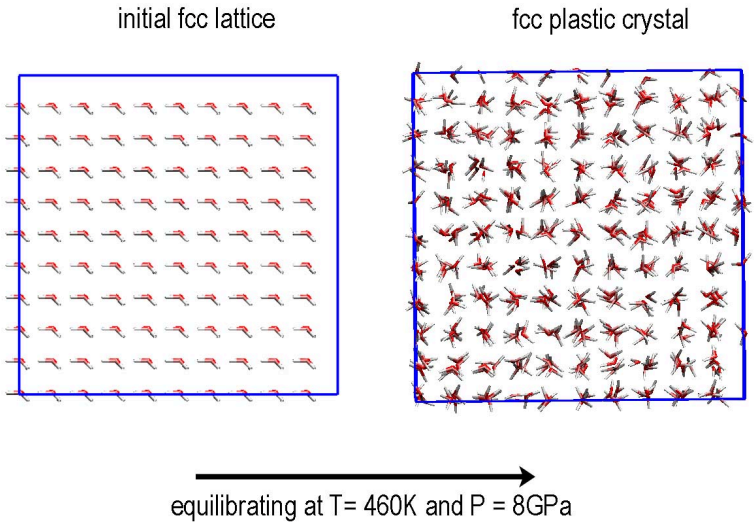

Figure 7.3: Diagram showing the initial configuration used for the fcc plastic crystal and the equilibrium configuration.

7.2.2 The oxygen-oxygen radial distribution functions

The structural radial distribution function were computed to evaluate quantitatively the results of the snapshots shown for the high pressure liquid, bcc and fcc plastic crystals and the ice vii crystal. Figure 7.4 shows the oxygen-oxygen $g_{OO}(r)$ vertically displaced from bottom to top for liquid water at ambient conditions (300 K, 100 kPa), liquid at 440 K and 8 GPa, bcc plastic crystal at 440 K and 8 GPa, fcc plastic crystal at 460 K and 8 GPa and the crystal ice vii at 340 K and 7 GPa respectively.

At the bottom of figure 7.4 the $g_{OO}(r)$ for liquid water at ambient conditions is shown. The high pressure liquid is shown displaced vertically, it has larger peaks corresponding to a high pressure liquid. The next phase shown is the bcc plastic crystal (red line): the $g_{OO}(r)$ is different to that of liquids by having stronger positional correlations seen as broad peaks. In a plastic crystal the broadening of the peaks occur because of the thermal agitation and its intrinsic orientational disorder. The fcc plastic crystal (green line) has the same chacracteristics of the bcc phase, although it has different peak position and heights which indicate a different positional lat-

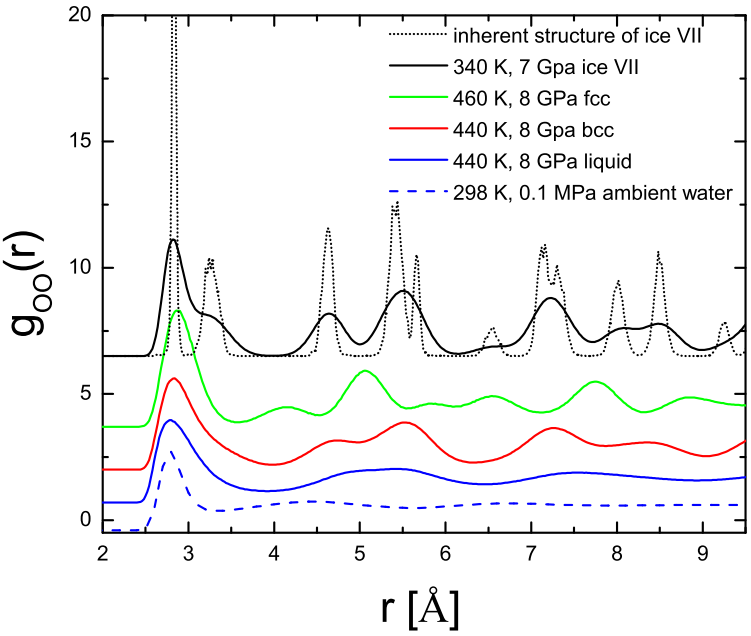

Figure 7.4: The oxygen-oxygen radial distribution function $g_{OO}(r)$. From bottom to top: liquid water, high pressure liquid, bcc and fcc plastic crystals and crystal ice vii. The inherent structure of ice vii is also shown.

tice. At the top of figure 7.4, the $g_{OO}(r)$ for ice vii is shown with a black continuous line: it is similar to those of plastic phases, but the peaks are higher. The broadening of the peaks correspond with the thermal agitation.

In order to see more clearly the positional ordering expected from the ice phase, the inherent structure was obtained for ten different configurations completely uncorrelated using an steepest descent method (4). The inherent structure contains the structure of the crystal without thermal vibrations. The resulting function is shown in figure 7.4 with a dashed black line: it shows sharp peaks on the position of the bcc lattice (the same as those of the bcc plastic). The inherent structure was also obtained for the fcc, bcc and liquid phases, but the results did not differ from their $g_{OO}(r)$ and are not shown. This confirms that the broadening of the peaks does not corresponds with thermal vibrations, but to intrinsic disorder in these cases.

7.3 Dynamics

In order to obtain more information on the different phases of water obtained in the previous section, their dynamics was analyzed using two functions: the mean square displacement (MSD) and the reorientational correlation function (RCF). The MSD displacement contains information about the diffusion of the molecules through the lattice, while the RCF has information about the rotational diffusion or how the molecules reorientate.

7.3.1 Mean square displacements

The MSD for all the phases is shown in figure 7.5. The MSD for the high pressure liquid shows that water molecules diffuse with a self-diffusion coefficient of $0.48x10^{-5}\,cm^2/s$, which is smaller compared to the value for TIP4P/2005 water at 298 K of $2.1x10^{-5}\,cm^2/s$. At the bottom of figure 7.5 three flat MSDs corresponding with the solid phases are shown: bcc and fcc plastic crystals and the ice vii crystal. There is no diffusion for this phases, there are only small displacements that occur because of the thermal agitation. The MSD in the plastic crystals is slightly higher than that of crystal ice vii and it is also higher in the bcc plastic compared to the fcc plastic phase. In order to better understand the dynamics of these phases the orientation is studied next.

7.3.2 Reorientational correlation functions

The relaxation properties for the orientation of the molecules can be characterized by the RCF: orientationally disordered phases have an RCF that decay to zero, while for an orientationally ordered crystal the orientations are fully correlated, being its value 1 at all time. The RCFs were computed as:

$$C_l(t) = \langle P_l[\mathbf{U}_\alpha(0) \cdot \mathbf{U}_\alpha(t)]\rangle \tag{7.1}$$

where P_l is the first (P_1) or second (P_2) rank Legendre polynomial:

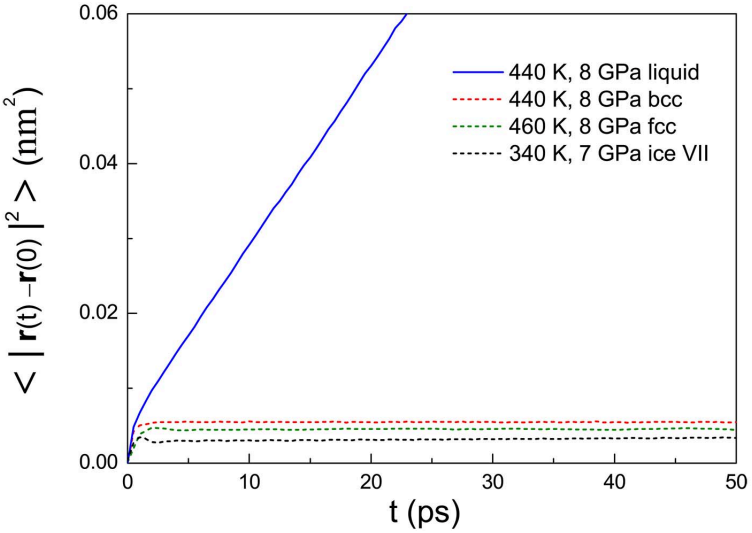

Figure 7.5: Mean square displacement. From bottom to top: ice vii, fcc and bcc plastic crystals with almost zero positional displacement and finally the high pressure liquid showing a diffusive motion.

$$P_1(x) = x \tag{7.2}$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1) \tag{7.3}$$

and $\mathbf{U}_\alpha$ is the unit vector which points along a given α axis in the water molecule. Four different axes were used; the O-H axis, the H-H axis, a perpendicular axis to the molecule defined as $\perp = \mathbf{r}_{OH1} \times \mathbf{r}_{OH2}$ and finally the molecular dipole μ. Some of these RCFs can be measured experimentally, for instance the H-H axis can be measured by $^1H\ -^1 H$ dipolar relaxation NMR experiments and μ can be related to dielectric relaxation measurements. Although the major problem is posed by the pressure range that is unaccessible nowadays.

Figure 7.6 shows the results for the RCFs of the μ vector of all the phases. The first Legendre polynomial is shown in the left side and the second Legendre polynomial is shown in the right side of the figure. The liquid and bcc plastic were simulated both at P = 8 GPa and T = 440K, but the

fcc plastic was obtained at T = 460 K. In order to have a direct comparison with the bcc and liquid phases, a simulation of the fcc plastic crystal at T = 440K was also performed. The RCF for the μ vector in the ice vii crystal remains close to 1: a full correlation meaning a perfect orientational crystal. However, the function for ice vii goes to a value close to 0.95, this means that there is a probability(0.05) for reorientations to happen. This reorientations account for the space group symmetry of the structure in the ice vii that is proton disordered and not to dynamical reorientations.

The RCFs for the liquid and plastic crystals share the same characteristics, they decay to zero showing that the molecules loose their orientational correlation after some characteristic time τ. Qualitatively τ can be inferred from the RCF: it will be smaller for fast decaying functions, in this case the plastic RCFs decay faster that the liquid one. That means that the molecules reorientate faster in the plastic crystal. This can be understood from the fact that the liquid has positional and orientational motions coupled, which is not the case of plastic crystals that use their energy mostly to rotate, having almost no displacements. From P_1 for the μ vector it can also be seen that the fcc crystal reorientates faster than its bcc counterpart. Notice that a simulation at the same temperature of the bcc phase has been added as a dotted line. Therefore, the faster reorientation in the fcc phase is a characteristic of the system and is not due to the thermal energy : the fcc lattice has 12 nearest neighbours, which makes it more compact and water molecules feel a stronger repulsion from its neighbours. The second Legendre polynomial P_2 decays faster, which is the expected behaviour.

In order to have a clear picture on the reorientation of the water molecule, the H-H, O-H and , μ and $\perp$ RCFs are shown in figure 7.7 for the liquid and plastic phases with the same colour code used in previous figures: blue for the liquid, red for the bcc plastic crystal and green for the fcc plastic crystal. The P_1 and P_2 functions are pictured in the same graph for each phase, with a clearer colour for the second Legendre polynomial. A comparison between the three phases shows that the molecules reorientate faster in the fcc plastic, decaying around 2 ps for all directions, the bcc plastic RCFs

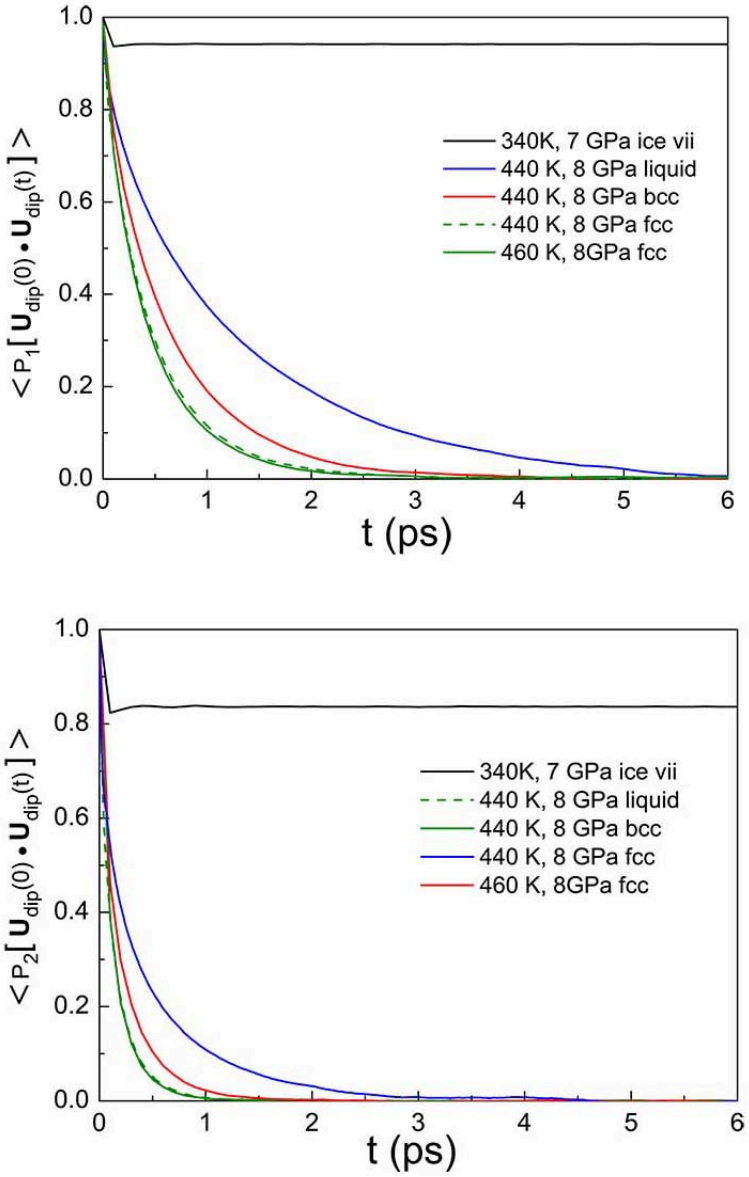

Figure 7.6: Reorientational correlation functions (RCF) for the dipole μ vector. At the top the first Legendre polynomial and at the bottom the second Legendre polynomial

decay around 3 ps and the decaying is slower in the liquid, around 5 ps. The rototranslational coupling in the liquid is the responsible of a slower decaying and also the higher packing in the fcc could cause a faster decay.

Direction	reorientation time	liquid	bcc	fcc
OH	$\tau_1^{\mathrm{OH}}(ps)$	1.13	0.66	0.46
	$\tau_2^{OH}(ps)$	0.60	0.24	0.16
	τ_1^{OH}/τ_2^{OH}	1.9	2.8	2.9
HH	$\tau_1^{HH}(ps)$	1.20	0.65	0.45
	$\tau_2^{HH}(ps)$	0.60	0.26	0.16
	τ_1^{HH}/τ_2^{HH}	2.0	2.5	2.8
perpendicular	$\tau_1^{\perp}(ps)$	0.80	0.51	0.38
	$\tau_2^{\perp}(ps)$	0.40	0.21	0.15
	$\tau_1^{\perp}/\tau_2^{\perp}$	2.0	2.4	2.5
dipole	$\tau_1^{\mu}(ps)$	1.20	0.63	0.45
	$\tau_2^{\mu}(ps)$	0.40	0.20	0.15
	$\tau_1^{\mu}/\tau_2^{\mu}$	3.0	3.1	3.0

Table 7.1: Reorientation times

The anisotropy of the reorientational dynamics can be seen by comparing the different axis RCFs, for instance, in the liquid RCFs it can be seen that the H-H, O-H and μ axis have similar reorientation times, but the $\perp$ axis reorientation is slower. This means that there is an anisotropy in the reorientation of water molecules in the high pressure liquid phase. This effect is smaller in the plastic phases, the $\perp$ axis has a closer reorientation decaying to those of the other axes. It means that the reorientation is more isotropic in the plastic phases. In order to make a better comparison, the numerical value of the relaxation time is needed.

The reorientation time τ can be obtained quantitatively in order to compare the different phases. It can be obtained by an integration of the RCF (7) and also by fitting the RCF according to the function $\exp^{-t/\tau}$ (9). Both methods were used and the relaxation times agree to the first decimal. The values obtained by fitting the exponential function up to 5 ps are shown in table 7.1.

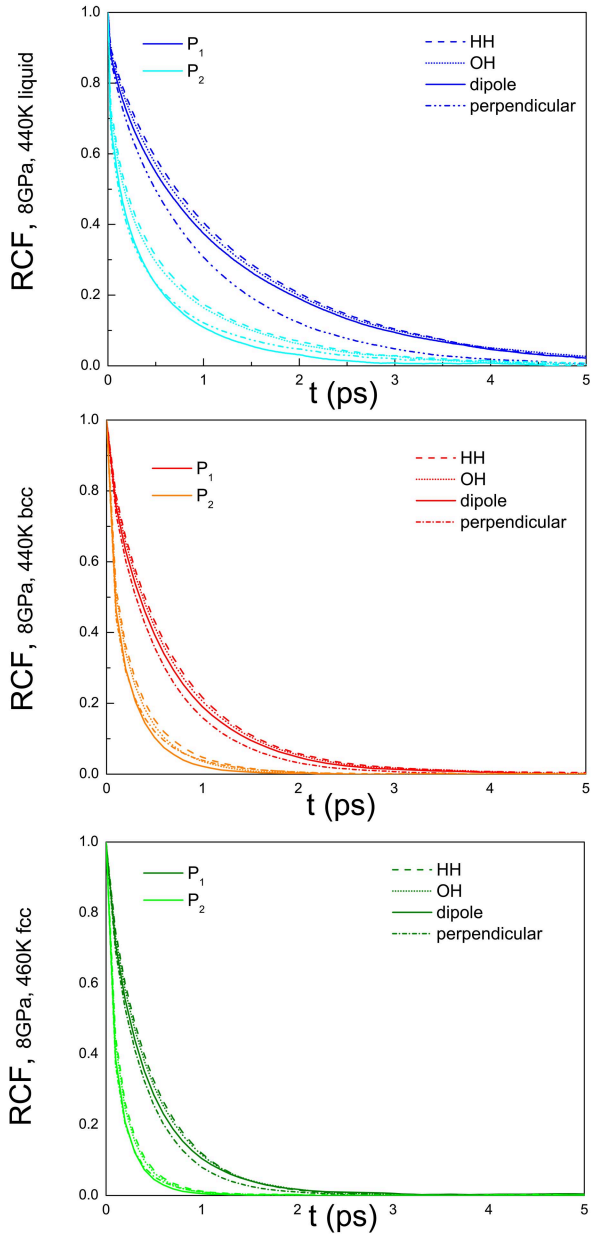

Figure 7.7: Reorientational correlation functions (RCF) for the OH, HH, dipole $\mu$and perpendicular $\perp$ vector. From bottom to top: fcc, bcc plastic crystals, high pressure liquid and crystal ice vii.

The ratio between the reorientation times τ_1 and τ_2 indicates the type of reorientation of the molecules. In the case of a rotational diffusion consisting of small angular jumps (almost continuously), the reorientation times decay exponentially with a time constant $\tau_l = [D_R l(l+1)]^{-1}$ (9) and then the ratio $\tau_1/\tau_2 = 3$. Where D_R is the so-called rotational diffusion coefficient and it can be measured experimentally. For rotations that consist of larger jumps the ratio τ_1/τ_2 is expected to be smaller.

The ratio τ_1/τ_2 is closer to the value of 3 in the plastic phases compared to the liquid phase. This is showing that the reorientation in the plastic phases occurs for smaller jumps compared to the liquid phase. In the plastic phases the ratios closer to 3 are the ones that correspond with the O-H and μ axis, which is an indication that the reorientation in these phases occur in an easier way around those vectors. On the contrary, the $\perp$ axis shows larger jumps according to the ratios closer to a value of 2 in all the phases. The perpendicular direction is not an easy direction for the reorientations. A final remark from the ratios is that they are closer to the value 3 for all the directions in the fcc phase, which agrees with the fact of a higher packing of the fcc plastic phase that favours an isotropic and faster reorientation of the water molecules.

7.4 Local ordering

7.4.1 Short range positional ordering

The results from the structure and dynamics in previous sections had indicated a different reorientation dynamics between the phases. In order to understand the influence of the structure in the short range on this behaviours the detailed positional distribution function $g(\theta_{\text{pos}}, \phi_{\text{pos}})$ was studied as in previous chapters.

The ice vii crystal and bcc plastic phases have both the same lattice, with 8 nearest neighbours. The fcc plastic phase has 12 nearest neighbours instead. According to this, the first 8 neighbours were analyzed for the liquid, bcc and ice vii phases and the 12 nearest neighbours in the case

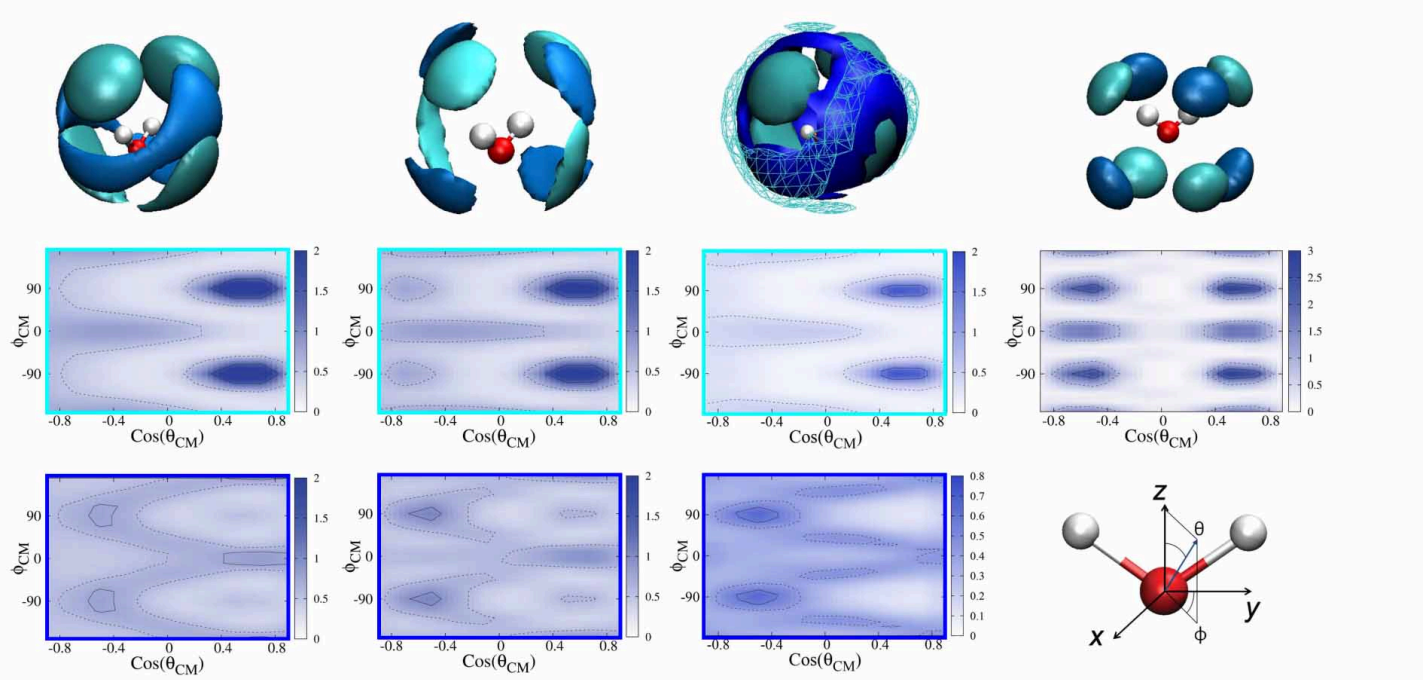

Figure 7.8: Spatial density maps (SDM) and distribution functions $g\left(\theta_{\mathrm{pos}}, \phi_{\mathrm{pos}} | r\right)$. From left to right: high pressure liquid, bcc and fcc plastic crystals and crystal ice vii. The upper distribution functions show neighbours 1-4 and the 8 first neighbours for ice vii (bcc unit cell). The lower distribution functions show the neighbours 5-8 in the liquid and bcc phases and the 5-12 neighbours for the fcc plastic crystal. First four neighbours are shown in cyan color and the next neighbours in blue. The inset shows the axis definition for the Angula analysis.

of the fcc plastic. Figure 7.8 shows the results for all the phases: the 1-4 neighbouring molecules at the top and the neighbours 5-8 at the bottom. In the case of the fcc plastic crystal, neighbours 5-12 are shown in the bottom panel. Only one map containing neighbours 1-8 for the ice vii is shown, this is because of its high symmetry compared to the plastic phases.The spatial density maps are shown along each map. This maps have been used before to characterize the oxygen distribution of neighbours around a central molecule. Soper and Ricci obtained the spatial density map for water at 0.4GPa using neutron diffraction and EPSR simulations (6), the pressure ranges studied in this thesis are yet unaccessible experimentally nowadays.

Analyzing the first 4 neighbours from left to right in figure 7.8 shows that the high pressure liquid has the same tetrahedral ordering as liquid water (cf chapter 5). The bcc and fcc plastic crystals show a similar positional ordering. Next, the 5-8 neighbours shown in the lower panels indicate also a tetrahedral ordering in antiphase to the first neighbours, this was something that Soper and Ricci noted from the experimental results at 0.4 GPa (6), although they were studying the liquid phase the same features are observed in the liquid at 8 GPa, the out of first shell neighbours are close to the first shell. The positional maps were studied also for a liquid at 6GPa and the same results were obtained (results not shown here). All the phases have in common two tetrahedrons in antiphase for their first 8 neighbours. Finally, The fcc phase has 4 additional neighbours that are located at $cos(\theta_{pos} = 0)$: the corners adjacent to a face with the reference molecule centered in that face.

The next question to address is: how similar the orientational ordering between these phases is?. In order to answer this, the orientational distribution functions were studied for the first eight neighbours in each phase and also separately for the 9-12 neighbours in the fcc plastic phase.

7.4.2 Short range orientational ordering

The orientational distribution functions $g\left(\theta_{\text{ori}}, \phi_{\text{ori}}\right)$ and $g\left(\theta_{\text{ori}}, \psi_{\text{ori}}\right)$ were obtained by selecting the regions in the positional maps corresponding to

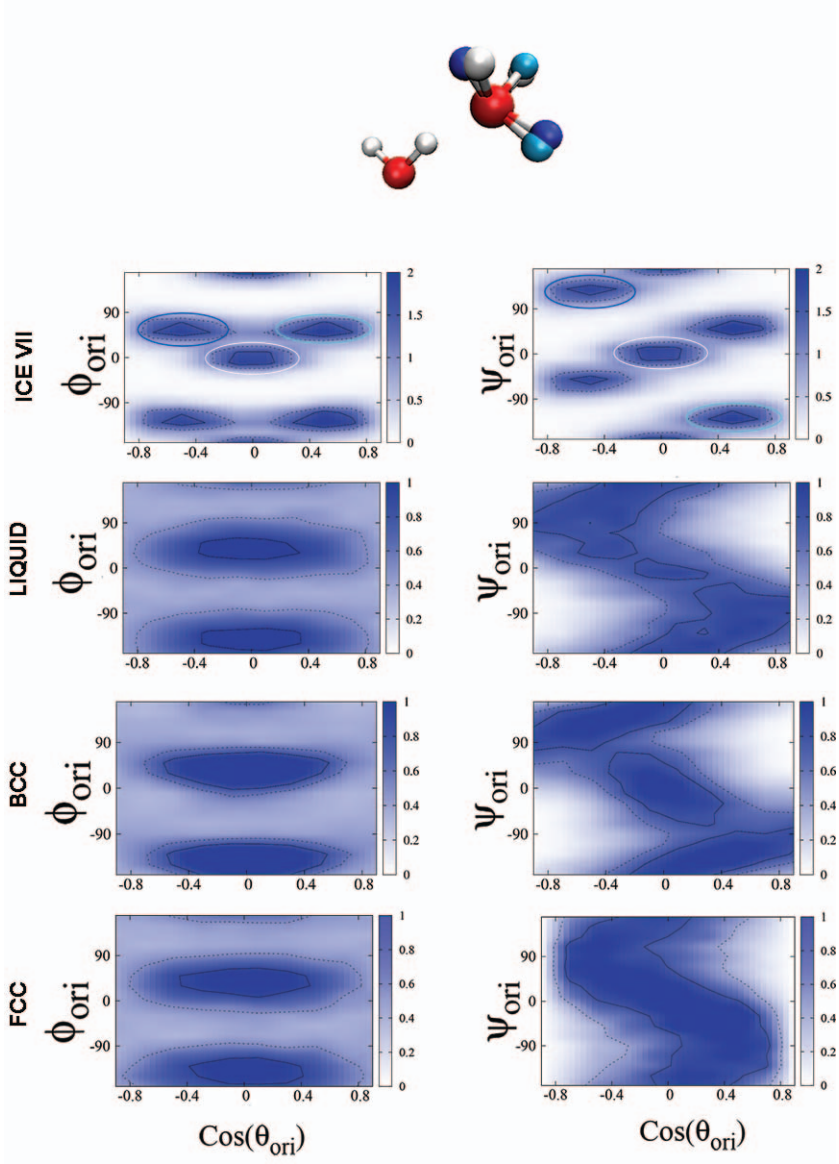

Figure 7.9: Left: Analysis of the orientational distribution functions $g(\theta_{\text{ori}}, \phi_{\text{ori}})$ and $g(\theta_{\text{ori}}, \psi_{\text{ori}})$ of the nearest neighbours in each phase. The phases are shown from bottom to top for the ice vii, liquid, bcc and fcc plastic phases. An inset at the top is shown with three of the most probable configurations.

neighbours 1-8 and the results are shown in figure 7.9. The analysis of the orientation for the 9-12 neighbours of the fcc plastic phase is shown sper-

ately in figure 7.10.

At the top figure 7.9 the orientational maps for the ice vii present well defined possible orientations. There is a correspondence between the high probability regions in both ϕ_{ori} and ϕ_{ori}. Three of those possible configurations were selected and highlighted with white, blue and cyan ellipses. The corresponding molecular orientations are shown as an inset at the top of the figure with the hydrogens with the same colours. The water molecule rotates while keeping its oxygen in the same position, this is hydrogen bonded to the reference molecule. A similar patter of possible orientation appears in the liquid, in this case the $g(\theta_{\text{ori}}, \phi_{\text{ori}})$ map shows two probable spots with more possible orientations, more in a continuum way contrary to well defined configurations in the crystal ice. When selecting the higher spot, corresponding to the spots chosen from the ice vii, a similar set of values was obtained for $g(\theta_{\text{ori}}, \psi_{\text{ori}})$. Again, there are more possible orientations, water molecules can pass between different orientations in a more continuous way.

By performing the same analysis on the plastic phases, the same $g(\theta_{\text{ori}}, \phi_{\text{ori}})$ maps were obtained. And in the $g(\theta_{\text{ori}}, \psi_{\text{ori}})$ maps an interesting result is that for the fcc plastic phase there is a continuous path between possible orientations. This agrees with the findings of the dynamic reorientation functions that show a diffusional rotation according to the ratios τ_1/τ_2 close to 3. On the one hand the fcc phase shows the faster reorientational dynamics and a continuum of possible orientations between the possible configurations of ice vii. On the other hand the bcc phase shows some not allowed orientations. This result has been shown by Aragones and Vega (1), a more uniform distribution of angles in the fcc phase: the molecules can jump between the barriers in the fcc while they can not do it in the bcc. Another fact from the fcc ordering is that it has 4 additional neighbours to the bcc plastic that are located coplanar to a molecule face centered (in the corners of the face).

In order to understand better the characteristics of the fcc plastic phase, the additional 9-12 neighbours were analyzed and the results are shown in

figure 7.10. When the coplanar neighbours were chosen, the $g(\theta_{\text{ori}}, \phi_{\text{ori}})$ map obtained shows that the possible orientations are any but those of the first 8 neighbours. This means that the coplanar molecules are highly orientationally disordered. This molecules can be the responsible for the faster reorientation in the fcc plastic phase.

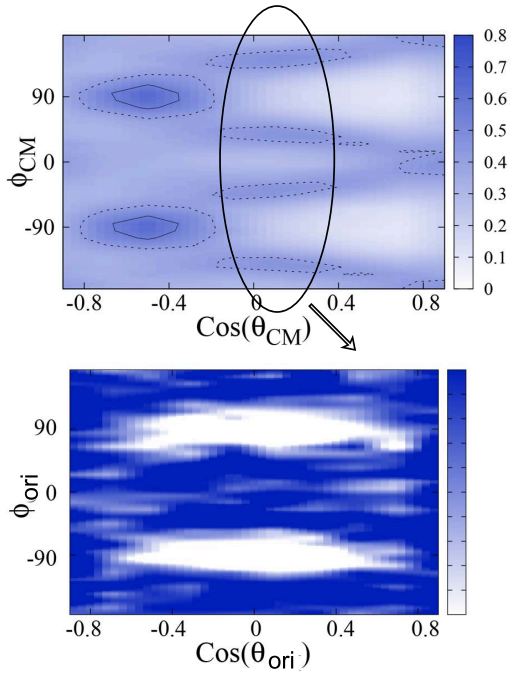

Figure 7.10: Top: Positional distribution function $g(\theta_{\text{pos}}, \phi_{\text{pos}})$ for the fcc plastic crystal. Bottom: Orientational distribution function $g(\theta_{\text{ori}}, \phi_{\text{ori}})$ of the molecules corresponding to neighbours within the same plane of a central molecule

7.5 Summary and conclusions

The plastic phases found by Aragones et al. and Takii et al. (1, 2) have been reproduced and the structure, dynamics and local ordering was determined.

The pressure range where the new plastic phases of water appear has not been reached yet experimentally. However, there are experiments in the Gigapascal range performed by Bove et al. (3) that study the transla-

tional and rotational diffusion up to 3 GPa. One of the conclusions of their work was that the pressure effect is more subtle in the structure of water compared to the temperature. This means that the high pressure does not change substantially the hydrogen bonding ordering, it only modifies the tetrahedral angles in a way that the compressed liquid coordinates the out of first shell neighbours towards a more compact structure. Soper and Ricci also found for water at 0.4 GPa (6) that water shows a first shell tetrahedron and a second tetrahedron in antiphase that gets closer with the pressure.

The results of this chapters show the presence of two interpenetrating tetrahedrons and the resemblance between the positional and orientational ordering between the high pressure liquid, plastic phases and ice vii crystal. The reorientation relaxation times show that the plastic phases present a near behaviour to the rotational diffusion model. The spatial density maps agree with the findings of Soper at 0.4 GPa, showing that the two tetrahedrons in antiphase are almost compacted to the same distance, this effect is higher for the pressure studied in this chapter of 8 GPa.

7.6 Bibliography

[1] J. Aragones and C. Vega. Plastic crystal phases of simple water models. *The Journal of chemical physics*, 130(24):244504, 2009. 115, 116, 117, 131, 132

[2] J. Aragones, M. Conde, E. Noya, and C. Vega. The phase diagram of water at high pressures as obtained by computer simulations of the tip4p/2005 model: the appearance of a plastic crystal phase. *Physical Chemistry Chemical Physics*, 11(3):543–555, 2009. 115, 116, 117, 118, 132

[3] L. Bove, S. Klotz, T. Strässle, M. Koza, J. Teixeira, and A. Saitta. Translational and rotational diffusion in water in the gigapascal range. *Physical review letters*, 111(18):185901, 2013. 132

[4] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Comput.*, 4:435–447, 2008. 120

[5] K. Himoto, M. Matsumoto, and H. Tanaka. Rotational dynamics of plastic ice. *Journal of the Physical Society of Japan*, 81(Suppl. A):SA023, 2012. 115

[6] A. K. Soper and M. A. Ricci. Structures of high-density and low-density water. *Physical review letters*, 84(13):2881, 2000. 129, 133

[7] I. Svishchev and P. Kusalik. Dynamics in liquid water, water-d2, and water-t2: a comparative simulation study. *The Journal of Physical Chemistry*, 98(3):728–733, 1994. 125

[8] Y. Takii, K. Koga, and H. Tanaka. A plastic phase of water from computer simulation. *The Journal of chemical physics*, 128(20):204501, 2008. 115, 117

[9] D. van der Spoel, P. J. van Maaren, and H. J. Berendsen. A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *The Journal of chemical physics*, 108(24):10220–10230, 1998. 125, 127

Chapter 8

Summary and conclusions

The understanding on the structure of disordered phases is crucial for many field like physics, chemistry, biology and technology. In this thesis molecular dynamics (MD) simulations of two disordered phases: the liquid and the plastic crystal were investigated for two different systems, hexachloroethane and water. The main contribution this work makes is about the use of new methodologies beyond the partial distribution functions to characterize the structure of these phases. Although generating the data accurately using MD was a major undertaking in itself, the real focus of the research was on understanding the dynamics and local ordering using an statistical method to find the probability distributions of the most probable position and orientation of molecules, information theory concepts and also a Bayesian approach.

In Chapter 4 an investigation of the structure and dynamics of the C_2Cl_6 liquid and bcc plastic crystal phases was done. The intra molecular structure of hexachloroethane was fitted to neutron diffraction experiments in the plastic phase giving an improved set of parameters for the plastic phase compared to previous electron diffraction experiments that were measured for lower temperatures. The liquid phase of hexachloroethane was obtained by MD in order to compare the results with the known plastic phase. The local ordering was studied by analyzing the probability distribution of positional and orientational variables and it was found that the positional and

orientational ordering is similar in the first neighbours. The long range positional order was also compared using information theory. The persistence of some positional ordering at distances beyond the first hydration shell was found for the liquid, something somehow unexpected.

In Chapter 5 an application of the positional and orientational probability distribution analysis was used in the case of liquid water at ambient conditions. The results presented in this chapter showed as in the work of Kühne *et al.* an asymmetry in the energy of water molecule contacts (whether it is the same as in that case must be further investigated). Nevertheless, the results demonstrated a clear difference of the geometry of these contacts thanks to a careful analysis of the local structure of liquid water. All of this suggests that a further analysis of the local structure of water must take this asymmetry into account which is not due to differences in the hydrogen bonding geometry but rather to different dipole orientations. The two most probable orientations found, give rise to two kinds of dimers: one close to the lowest energy of a water dimer in vacuum with an almost perpendicular alignment of the dipole moment, and another one with a parallel orientation of the dipole moment which is less tightly bound. These two different dimers have an effect on the orientation of further water dipole moments up to a distance of ≈6Å. Liquid water can therefore be described as a continuous mixture of two kinds of dimers where the Hydrogen bonds have the same geometry but the interaction energies are different due to a different mutual orientation of the dipoles of the participating water molecules.

In Chapter 6 the information theory was used to compute the excess entropy calculations to determine if changes in molecular ordering and in local density are correlated with the relative position or orientation of molecules. Using these functions it seems that the ordering of liquid water at distances larger than ∼4 Å is ruled for distances up to about ∼6 Å by an alternating *orientation* of water dipoles, while at distances between ∼6 Å and ∼9 Å it is the *position* of water molecules that dominates the changes in molecular ordering. Even at distances as long as 9 Å there are "shells" which reproduce the well-known tetrahedrality of the first four water molecules. This

special positional order could constitute a new mechanism for long-distance interactions between molecules in aqueous solution, such as for example in molecular recognition or protein-protein interactions, which would play a role together with long dipole-chains reported in the literature.

In Chapter 7 an investigation of the new plastic phases of water found by Aragones et al. and Takii et al. was performed. The dynamics and local ordering was determined for those phases and for a high pressure liquid and ice vii. The results of this chapters showed the presence of two interpenetrating tetrahedrons and the resemblance between the positional and orientational ordering between the high pressure liquid, plastic phases and ice vii crystal. The reorientation relaxation times showed that the plastic phases present a near behaviour to the rotational diffusion model. The spatial density maps agreed with the findings of Soper at 0.4 GPa, showing that the two tetrahedrons in antiphase are almost compacted to the same distance, this effect was higher for the pressure studied in this chapter of 8 GPa.

Taken as a whole, the research presented in Chapters 4-7 of this thesis represent a small contribution to the understanding of disordered phases. There is a vast amount of systems that can be studied using the methodologies in this work. A particular promising field is that of biochemistry. With the increasing power of computers, it is now possible to unravel problems as fundamental to life as that of folding of proteins. The interaction of biomolecules with water and other substances at the atomic scale is key in the understanding of biological processes. The methods presented in this thesis can provide with powerful insights in the atomic structure and interactions.

8.1 Conclusions

The general conclusions that can be drawn from this thesis are divided in the methodological and scientific contributions:

8.1.1 Methodological contributions

- A method to find the most statistical relevant position and orientation of molecules of disordered phases have been developed.
- Information theory concepts have been applied to the study of the local and global ordering of molecules in disordered systems.
- A Bayesian methodology have been implemented to obtain the intra molecular structure from experimental data.

8.1.2 Scientific contributions

- It has been found that the structure of a simple liquid as C_2Cl_6, for which the steric effects are the most important to describe the structure, the liquid and plastic crystal phases are similar in the short range positional and orientational ordering and that the resemblance also persists in the long range positional ordering.
- It has been found that the local ordering in liquid water in the short range order is the result of a continuous mixture of molecules in two kind of dimers: on the one hand *trans* dimers with almost perpendicular dipole moments and a low energy and on the other hand *cis* dimers with parallel dipole moments and a different energy, being less tightly bound.
- Two regimes for liquid water have been found beyond the first hydration shell using excess entropy calculations: a regime ruled by the alternating *orientation* of water dipoles at distances larger than $\approx 4\,\mathring{A}$ up to about $\approx 6\,\mathring{A}$ and a second regime between $\approx 6\,\mathring{A}$ and $\approx 10\,\mathring{A}$ distances where the *position* of water molecules dominate the changes in molecular ordering.
- It has been also found that for a non simple liquid, as it is the case of water, that the liquid and plastic crystal phases have a similar local positional and orientational ordering.